Papers in Southeast Asian linguistics No.15

Further Chamic Studies

PACIFIC LINGUISTICS

PAPERS IN SOUTHEAST ASIAN LINGUISTICS NO.15
FURTHER CHAMIC STUDIES

edited by
David Thomas

TABLE OF CONTENTS

FOREWORD

This volume of studies on Chamic languages comes two decades after the previous volume (*Papers in Southeast Asian Linguistics* No.4: *Chamic Studies*, PL, A-48, 1977), and so seems long overdue in a field of study that is active. The first volume included only papers on languages in or near the Champa homeland; this second volume looks farther afield, with a paper on Tsat, a clearly Chamic language in Hainan, and with glances towards possible further affiliation with Acehnese and Moken. With the recent renewed scholarly interest in Chamic we can hope that the third volume of studies on Chamic languages will not be too long in coming.

David Thomas

Summer Institute of Linguistics

A GRAMMAR SKETCH OF WESTERN (CAMBODIAN) CHAM

NEIL I. BAUMGARTNER

1. INTRODUCTION[1]

The Cham language is in the Chamic branch of the Austronesian family of languages. It is spoken by about 300,000 to 350,000 people in Vietnam and Cambodia. The language is divided into two major dialects, or more likely, two separate languages. Eastern (or Vietnamese) Cham is spoken by about 35,000 people in Vietnam in the area of the towns of Phan Rang and Phan Ri, and Western (or Cambodian) Cham (WCham) is spoken in Cambodia by about 250,000 to 300,000 people and by about 20,000 people in the Mekong Delta region of Vietnam, with concentrations in Chau Doc, Tay Ninh, and Saigon. There are three main divisions of Western Cham—Western Cham as spoken along the Mekong, Western Cham as spoken along the Tonle Sap especially in the Kompong Chhnang province, and Western Cham as spoken in Vietnam—as well as a fourth small group of Chams near Kompot. Pronunciation accounts for much of the difference between the dialects, with vocabulary also being a factor. There seem to be few, if any, differences in grammatical structure.

This is a grammar sketch of Western Cham and is based on texts that were collected in 1970–1975 in Southeast Asia by Timothy and Barbara Friberg, who were working there with the Summer Institute of Linguistics. The texts represent several of the dialects of Western Cham. These texts have been supplemented by a small amount of elicited data. Many questions still exist and may be answered after looking at more data.

2. CLAUSES

The basic clause structure of WCham is subject–verb–object (SVO). Adverbial elements—words or phrases that modify a verb or an entire sentence—occur at four places in the clause. They may occur at the beginning of the clause (conjunctions), at the end of the clause (final particles), before the verb (preverbal adverbs) or after the verb (adverbs). Embedded adverbial clauses may occur at either the beginning or end of the clause.

2.1 BASIC CLAUSES

Basic clauses in WCham are of two types, active and non-active (also known as stative).

---

[1]   See the appendix for a list of abbreviations.

David Thomas, ed., *Papers in Southeast Asian linguistics* No.15: *Further Chamic Studies*, 1-20.
Pacific Linguistics, A-89, 1998.
© Neil Baumgartner

1

2.1.1 ACTIVE CLAUSES

2.1.1.1 AMBIENT ACTIVE CLAUSES

Only one example of an ambient clause has been found so far. It contains the word *djauk* to hit' with the phrase *ngĭn rabuk* 'storm':

(1)   *Djauk   ngĭn   rabuk.*[2]
         VT   N   <
         to.hit storm   <
         There arose a storm.

More examples are needed before a description of the structure of ambient clauses can be given.

2.1.1.2 INTRANSITIVE ACTIVE CLAUSES

Intransitive clauses have the structure of noun phrase as subject followed by the verb phrase. For example:

(2)   *Ku   nơn   đuaik   nao.*
         N   DET   VI   DIR
         Khmer that   to.run away
         The Khmers run away.

(3)   *Nhu   đuaik   nao.*
         PRO   VI   DIR
         3   to.run away
         They all ran away.

2.1.1.3 TRANSITIVE ACTIVE CLAUSES

Transitive clauses have the structure of noun phrase as subject, verb phrase, and noun phrase as direct object. For example:

(4)   *Hlŭn   mayai   ha   rŭng.*
         PRO   VT   NUM   N
         1SG.LORESP   to.say one   story
         I tell a story.

2.1.1.4 BITRANSITIVE ACTIVE CLAUSES

Bitransitive clauses have the subject noun phrase followed by the verb phrase with the direct object noun phrase followed by the indirect object prepositional phrase. The prepositional phrase is introduced by the preposition *ka* 'to'. For example:

---

2   The orthography used here follows Vietnamese conventions.

(5)   *Yah   rôk   nao,   ƀôh   drăp   kau,*
      ADVZ   VT   DIR   VT   N   PRO
      if   to.dig   away   to.see   thing   1.RESP

   hư   djauk ta   bray   ni   ka   kau   wơk.
   PRO   AUX   ADV   VB   DET   PREP   PRO   FIN
   2.LORESP must   only   to.give this to   1.RESP back
   If, when digging, I find my belongings, you must give this territory back to
   me.
   (*Rôk* means literally 'to dig around something that can be seen'.)

2.1.2 NON-ACTIVE CLAUSES

Non-active clauses are clauses that do not show any action in the verb. This type of clause includes equative clauses, descriptive clauses, adverbial clauses (such as 'he is here'; 'he is at home') and existential clauses.

2.1.2.1 EQUATIVE NON-ACTIVE CLAUSES

Equative non-active clauses contain two noun phrases. The first (in bold) functions as the subject of the clause and the second (underlined) is the predicate complement. There is no verb. For example:

(6)   **Kau**   <u>kra</u>   <u>kaɗuh.</u>
      PRO   <u>N</u>   <u>≤</u>
      1.HIRESP <u>turtle</u> <u>≤</u>
      I am a turtle.

2.1.2.2 DESCRIPTIVE NON-ACTIVE CLAUSES

Descriptive clauses have the structure of noun phrase as subject and adjective phrase as the predicate complement. In my data there are no words that function as a verb in descriptive clauses. For example:

(7)   *Nhu   uan   tabuan.*
      PRO   ADJ   <
      3   happy   <
      They (the Khmers) were very happy.

(8)   *Nhu   uan   tabuan   saƀai   tai   lô.*
      PRO   ADJ   <   ADJ   <   ADV
      3   happy   <   happy   <   very
      They were very, very happy.

2.1.2.3 EXISTENTIAL NON-ACTIVE CLAUSES

An existential clause in WCham may have the structure: verb phrase, consisting of the existential verb *mada* 'there is', followed by a noun phrase. For example:

(9)   *Mada   Chăm   ha   rang   nơn   trah.*
     VN   N.PROP   NUM   CL   DET   VI
     there.be Cham   one   person that to.cast(fish-nets)
     There was a Cham man casting fish-nets.

2.1.3 ADVERBIAL ELEMENTS

Adverbial types of elements can occur in four positions in a clause, but generally a particular adverbial word will be found in only one of these positions. Adverbial elements can occur at the beginning of a clause, at the end of a clause, before a verb or after a verb. Those that occur sentence-initially (examples 10,11) generally carry a meaning of time or of sequencing of events in relation to other events and relate the entire sentence to a clause or sentence preceding it. Typically, these are called conjunctions. They differ from conjunctions that connect two clauses into sentences or two phrases or words. They seem to function at the discourse level and tie the time-line of the story or episode together. They also differ from adverbialisers (i.e. subordinating conjunctions) in that adverbialisers connect a subordinate clause to its main ·clause, with the subordinate clause providing background or setting information, while the main clause usually presents new information. This type of conjunction will be called a sentence conjunction (CNJS).

(10)   **Bloh**   *patao   Chăm   laik:*
     CNJS   N   N.PROP   VT
     then   king   Cham   to.say
     Then the Cham king said:

(11)   **Hani**   *kau   bray   hư   hu   nưk   ha   rang.*
     CNJS   PRO   VT   PRO   VT   N   NUM   CL
     now   1.HIRESP to.give 2.LORESP to.have child   one   CL (person)
     Now, I will give you a child.

Adverbial elements that occur at the end of the clause or sentence have the entire clause or sentence in view and semantically carry the idea of completion (finished, completed, already), certainty (indeed, true), negation, possibility or impossibility, immediateness, or customary action (often, again). For lack of a better term these will be called sentence-final particles (FIN). For example:

(12)   *Ru   **bloh**   nưk   nơn   matai   **yơ**   dok   kađong.*
     VT   **FIN**   N   DET   VT   **FIN**   VI   ADV
     to.rock   **finished** child that to.die **already** to.stay quite
     Finished rocking, the child was dead and silent.

Preverbal adverbial words (PVA) carry the meaning of 'ever, nearly, only, still, always, or again'. These are part of the verb phrase and occur after the tense or auxiliary and before the verb. For example:

(13)   *Miưk   **đêl**   ngăk   sang   ray?*
     N   **PVA**   VT   N   YN.QM
     younger.uncle **ever**   to.make house   also
     Have you (young uncle) ever built a house?

The postverbal adverbial words (ADV) are what are more typically thought of as adverbs and usually relate the manner of action (melodiously, clumsily, well, much, for pleasure) or location or position (in a line, far). For example:

(14) *Ông   nơn   hamĭt   nhu mayai bangi   păng.*
      N   DET   VT   PRO VT   **ADV**   <
      TITLE.RESP   that   to.hear   3   to.talk **melodious** <
      He heard them speaking melodiously.

Of course, there are exceptions to the above, both in that some words occur in more than one place, and some places occasionally have a meaning other than what is listed above.

2.2 CLAUSE VARIATIONS

2.2.1 NEGATION

Negative clauses, as in examples (15) and (20), are formed by adding the final particle *ô* 'NEG' at the end of the clause.

(15) *Rean   ngăk pap gah   nŭk   matau   nơn ô.*
      VT   VT   ADJ N   N   N   DET **NEG**
      to.dare to.do evil direction child child.in.law that **NEG**
      He didn't dare do any more evil things to the son-in-law.

2.2.2 CHANGES IN GRAMMATICAL RELATIONS

In Cham, grammatical relations are indicated almost entirely by word order as there are no inflectional affixes, case markings or agreement markings. Passivisation has been found, but as yet no dative movement or other similar changes have been observed.

Passive clauses in Cham, as in examples (17) and (18), are formed by moving the noun phrase functioning as direct object to the beginning of the sentence and inserting *djauk* 'PASS', a passive marker or auxiliary verb, between the direct object noun phrase and the subject noun phrase.

(16) *Sohput khan   hlŭn.*
      N   VT   PRO
      friend to.tell 1SG.LORESP
      A friend told me.

(17) *Hlŭn   djauk sohput khan.*
      PRO   AUX   N   VT
      1SG.LORESP PASS   friend to.tell
      I was told by a friend.

An impersonal passive can be formed by using *rang* 'someone' after *djauk*. Whether the subject noun phrase can be omitted completely has not yet been determined.

(18)   *Hlŭn   djauk rang   khan.*
      PRO   AUX   PRO   VT
      1SG.RESP   PASS   3.INDEF   to.tell
      I was told by a friend.

2.3 EMBEDDED CLAUSES

Embedded clauses are clauses that are contained as part of the main clause, such as a relative clause, a complement clause functioning as the subject or object of a clause, or a clause that replaces an adverb phrase showing time, location, purpose, reason, and so on. Constructions such as indirect quotes, embedded questions, embedded commands, and subject-to-object raising have not yet been thoroughly analysed.

2.3.1 RELATIVE CLAUSES

A relative clause, as in examples (19)–(21), is a clause that is embedded in a noun phrase and that modifies the head noun of the noun phrase. Relative clauses are positioned just before the determiner (DET) if one is present, or at the end of the noun phrase if there is no determiner. The phrase in the relative clause that is the same entity as the head noun is replaced by *kung* 'who/what/which/when, etc.', the relative pronoun in Cham. In the following example, the relative clause is a non-active descriptive clause. The relativised noun phrase in the relative clause functions as the subject. Note that in these relative clause examples the head noun that is modified by the relative clause is underlined and the relative clause is in bold.

(19)   *No,*   *tăl   hray ha   sa   nơn mada   dăm   ha   rang*
      PART   ADVZ   N   NUM   <   DET VN   <u>N</u>   NUM   CL
      PAUSAL.PART   when   day one   <   that there.be <u>suitor</u> one   CL

      kung   **nas,**   *mai   duh   daok   ông   nơn.*
      REL.PRO **ADJ**   VI   VT   <   N   DET
      who   **intelligent** to.come to.serve <   old.man that
      One day it happened that an intelligent suitor came to serve the father.

In the following example, *dăm* 'suitor' is the head noun of the noun phrase that contains the relative clause. The relativised noun phrase is the subject and is realised by the relative pronoun *kung* 'who'.

(20)   *Yau*   *nơn yơ* <u>*dăm*</u>   *tŏng hadôm* **kung**   *mai*   *mơng*
      CNJS   <   <   <u>N</u>   ADJ ADJ   **REL.PRO**   VI   PREP
      that's.why   <   <   <u>suitors</u> all many   **who**   to.come from

      dahlau   **mai**   *nơn hu*   *drơm   hơng ông   nơn*
      N   **FIN**   DET VT   VI   PREP N   DET
      before   **toward**   that to.be.able.to to.endure with old.man that

      ngăk   *pap ô.*
      VT   ADJ FIN
      to.make   evil NEG
      Therefore all those previous suitors could not endure the man's evil deeds.

Relative clauses have been found in noun phrases that are a subject, direct object, predicate complement (predicate in non-active clauses) or a vocative or addressive. The relativised phrase in the relative clause can be a noun phrase, or adverbial phrase. The relativised phrase, whether a noun phrase, adverb phrase, or predicate complement, is replaced by *kung*. If the relativised phrase is not the subject of the relative clause, *kung* will be moved to the front of the clause. Below is an example of the relativised phrase being an adverb phrase. The noun phrase containing the relative clause is underlined, and the relative pronoun, *kung*, is in bold.

(21) *Tăl   hamĭt yau nơn nao   dăm   nơn nao blay kan mơk*
      CNJS VT   PRO <   FIN   N   DET VI VT   N   VT
      when to.hear thus <   away suitor that to.go to.buy fish to.take

      nao   kăk   dalăm ea   lăm   bơng kung   ông   nơn
      DIR   VT   PREP N   PREP N   REL.PRO   N   DET
      away to.tie.up in   water in   lake where   old.man   that

      khea   mơk   nao   trah   pagê.
      VT   VT   VI   VT   N
      to.propose to.have to.go to.fish.with.net tomorrow
      When the suitor heard this he went and bought fish and took them to put them in the lake where the old man planned to fish the next day.

2.3.2 COMPLEMENT CLAUSES

Complement clauses are clauses that are used in place of a subject, object, or some other nuclear element of the clause. (Subject complements have not yet been observed in Western Cham.) Direct and indirect quotes are also complement clauses. The only complementiser found so far is *laik* 'that'. Not all complement clauses use a complementiser.

2.3.2.1 OBJECT COMPLEMENTS

In Cham, with certain verbs, the direct object of a clause (i.e. the noun phrase that follows a verb in a transitive clause) may be a full sentence (i.e. an object complement). In example (22) the subject of both the main clause and the embedded clause is *Jawa*, so *Jawa* does not appear in the embedded clause. In (23) the subject of the main clause (Khmer, understood from context) and the subject of the embedded clause are different, so the subject of the embedded clause (*Malayu*) must appear. The complement clauses in (22) and (23) do not use a complementiser to introduce them.

(22) *Jawa   khĭn   bơng kŏk   ta-uk.*
      N   VT   **VT**   **N**   <
      Javanese to.want **to.eat knee** <
      The Javanese want to eat their knees.

(23) *Bôh   Malayu   mai   gaik.*
      VT   N   VI   ADV
      to.see **Malay**   **to.come again**
      The Khmers see the Malays coming again.

Another type of complement clause uses a complementiser to introduce the complement clause. The main clause verb is a verb of mental activity (e.g. think, want,[3] wonder, believe, hope, know, understand) as in examples (24) and (25), or a verb of speech (e.g. say, ask, cry, scream) as in (26). The clauses using verbs of speech will be presented in the section on direct quotes. The complementiser, *laik* 'that' is in bold, and the embedded clause is underlined.

(24) *Nhu thau*   **laik**   *Patao* *Chăm*   *nơn*   *tăk*   *phŭn* *krĕk* *nơn* *pajalơh*.
      PRO VT   **CMPZ** N   N.PROP DET   VT   N   N   DET   VT
      3   to.know **that** King   Cham   that   to.cut tree   krek   that   to.destroy
      They knew that the Cham king had cut the krek tree down.

(25) *Dray* *yŏl*   **laik**   *ngăk* *yau*   *nơn* *jiang* *ô*.
      PRO   VT   **CMPZ**   VT   PRO   <   NEG   <
      1PL   to.understand **that**   to.do   like.that   <   can't   <
      We understand that we cannot do like that.

2.3.2.2 DIRECT QUOTES

In a clause in which the main verb is a speech act, the direct object is an embedded clause, which is preceded by the complementiser *laik* 'that'. For example:

(26) *Ông*   *Chăm*   *nơn*   *sua*   **laik:**   *"Djauk*   *gêk* *phông*
      N   N.PROP   DET   VT   **CMPZ**   WH.QM   <   N
      TITLE.RESP   Cham   that   to.ask   **that**   why   <   group

      hư   *đuaik?"*
      PRO   VI
      2.LORESP   to.run
      An old Cham man asks them: "Why are all of you running?"

2.3.3 ADVERBIAL CLAUSES

Adverbial clauses are clauses which are embedded in another clause and which give information about the time, purpose or reason of the action of the clause in which it is embedded. An adverbial clause may also give the hypothetical or contrafactual condition (i.e. if…then) for the clause in which it is embedded. This is not the conjoining of two clauses, but rather the embedding of one clause inside another, with the embedded clause taking the place of an adverbial phrase. The embedded clause is introduced with an adverbialiser (ADVZ) which some may call a subordinating conjunction. In the following examples the adverbialiser is in bold, and the embedded clause is underlined.

2.3.3.1 TIME ADVERBIAL CLAUSES

Embedded time Adverbial Clauses usually come at the beginning of the sentence, with the main clause following it. For example:

---

3   The verb *khĭn* 'to want' is found both with and without a complementiser. See example (22).

(27) **Haday** **mơng** *thau*   *brŭk*   *nhu* *bloh,*   Patao Chăm
      ADVZ <   VT   N   PRO   FIN   N   N.PROP
      after <   to.know   matters 3   already   King   Cham

      ko   *bray* *nhu* *dok.*
      FOC   VT   PRO   VI
      FOC   to.let 3   to.stay
      When he knew their situation the Cham king let them settle.

2.3.3.2 PURPOSE ADVERBIAL CLAUSES

To indicate purpose, the purpose clause is embedded at the end of the main clause and it is introduced by the adverbialiser *tôk* 'in order that'. For example:

(28) *Dray* *ngăk*   *saphơu* *ni*   *yau*   *ni* *tôk*   *nŭk*   *neh* *nhu*
      PRO   VT   N   ADJ   ADV   <   **ADVZ**   N   <   PRO
      1PL   to.make   book   this   like.this <   **in.order.that** children <   3

      takŭ   *tai* *hi*   *baik.*
      VT   <   TNS   VT
      to.like <   FUT   to.study
      We make the book like this so that the students will want to study.

2.3.3.3 REASON ADVERBIAL CLAUSES

Reason is indicated by embedding a clause at the end of the main clause, introducing the embedded clause with the adverbialiser *kayoa* 'because'. The second clause gives the reason for the first. For example:

(29) *Nhu* *thau*   *laik*   *yau*   *nơn* *bray* *mai*   *nơn* **kayoa**
      PRO   VT   VT   COMP.ADV <   VT   DIR   DET   **ADVZ**
      3   to.know to.say   like.that <   to.let toward   that   **because**

      mai   *hi*   *pajalơh*   *phŭn* *krĕk.*
      VI   TNS   VT   N   N
      to.come FUT   to.destroy tree   krek
      They say that because she came to destroy the 'krek' tree.

2.3.3.4 HYPOTHETICAL CONDITION ADVERBIAL CLAUSES

Hypothetical condition clauses are if…then clauses. The embedded conditional clause (the 'if' clause) is first and is introduced by *yah* 'if', with the main clause giving the conclusion. For example:

(30) **Yah** *phŭn* *kayau* *nơn* *prung,* nao jhŭl, nao pajalơh
    ADVZ N < DET ADJ VI VT DIR VT
    if trees < those big to.go to.push away to.destroy

phŭn *kayau* *nơn.*
N < DET
trees < those
If the tree was big it would push it over.

2.3.3.5 CONTRAFACTUAL CONDITION ADVERBIAL CLAUSES

A contrafactual condition clause is an if...then clause which could be stated as 'if this were true, but it isn't...'. Contrafactual clauses in Western Cham have the same structure as hypothetical clauses. The identification of them as contrafactual comes from the context. In the following example, the one who came from India had asked permission to live there.

(31) **Yah** *hư* *thau* *laik* *tanưh* *ea* *ni* *drăp* *hư*
    ADVZ PRO VT CMPZ N < DET N PRO
    if 2.LORESP to.know that territory < this thing 2.RESP

nơn, *hư* *mai* *mơng* play *India* *dơh* *mai,*
DET PRO VI PREP N N.PROP PREP DIR
that 2.RESP to.come from country India over.there toward

hêt *gêk* *hư* *mai* *lakau* *di* *kau* *dok?*
N WH.QM PRO VI VT PREP PRO VI
reason why? 2.RESP to.come to.ask from 1.RESP to.live
If you knew that this territory was yours, when you came from India why did you ask us permission to live here?

2.4 CLAUSE COMBINING

Two or more clauses can be combined either with a Conjunction (CONJ) or by simply placing one clause after the other. With the tendency to leave out known information, the subject of a juxtaposed second clause can be omitted, especially if it is also the subject of the first clause.

3. PHRASES

3.1 NOUN PHRASES

3.1.1 HEAD AND MODIFIERS

In Western Cham, the head of the noun phrase (NP) can be a noun (N) (examples 37–41), pronoun (PRO) (examples 31, 33), a proper noun (N.PROP) (example 32), a determiner (DET) (example (34)) or classifier (CL) (examples 35, 36). A pronoun, proper noun, or determiner when used as the head of a noun phrase will always appear alone.

(32) *Ai*   *long lakay,*   *ai*   *long kamay,*
      N   <   ADJ   N   <   ADJ
      oldest.sibling   <   male(human)   oldest.sibling   <   female

      mai   păng   Idares   mayai takal ha   ruּng gaik.
      VI   VT   N.PROP   VT   N   NUM   N   ADJ
      to.come to.listen Idris   to.tell story one   story more
      Brothers and sisters, please come close to hear Idris tell another story.

(33) *Hlŭn*   *lakau   ma-ah di   doa   rang diuk pasang*
      PRO   VI   <   PREP NUM   CL   N   N
      1SG.RESP   to.pardon <   from two   CL   wife husband

      ai   long.
      N   <
      oldest.sibling   <
      I would like to ask the two of you (the couple) for pardon.

(34) *Ni mayai   mơng ruּng tanuּh   ea Chăm   mơng asăl   awăl*
      DET VT   PREP N   N   <   N.PROP PREP N   <
      this to.speak about history kingdom <   Cham   from origin <

      beah   mai   tăl   Chăm   bih   tanuּh   ea.
      ADVZ   DIR   PREP N.PROP   VT   N   <
      until   towards at   Cham   completely   kingdom   <
      This is to speak about the Cham kingdom, from the beginning to the complete
      disappearance of it.

When used as the head of the noun phrase, a classifier will be preceded by at least one
number (NUM). For example:

(35) *Tajuh   tapăn   rang*   *nao   tăl kuּh   glai   nơn.*
      NUM   NUM   CL   VI   <   ADV   N   DET
      seven   eight   person to.arrive.at <   middle forest that
      Seven or eight people went to the middle of the forest.

(36) *Jawa   blay   tangơi di   Chăm   ha   ratuh   rial.*
      N   VT   N   PREP N.PROP   NUM   NUM   CL
      Javanese to.buy corn   from Cham   one   hundred riel
      The Javanese people buy corn from the Cham for a hundred riels.

If a noun is the head of the noun phrase, it can be followed by an optional possessive NP
(examples 37, 38), an adjective (ADJ) (examples 37, 38) (the order of the NP and ADJ can be
switched), a number (NUM) (examples 37–39) and classifier (CL—the number cannot appear
without the classifier) and finally by a determiner (example 39).

(37) *Nhu bôh   sang   hlŭn   prung klau   bôh.*
      PRO VT   N   PRO   ADJ   NUM   CL
      3   to.see house 1SG.RESP big   three CL
      He saw my three big houses.

(38)   Nhu   bôh   <u>sang</u>   <u>prung</u> <u>hlŭn</u>   klau   bôh.
       PRO   VT   **N**   **ADJ**   **PRO**   NUM   CL
       3   to.see house big   1SG.RESP   three CL
       He saw my three big houses.

(39)   Đuaik   nao   tăl   <u>labik</u>   <u>ha</u>   <u>sa</u>   nơn,   ông   Chăm
       VI   DIR   PREP   N   NUM   CL   DET   N   N.PROP
       to.run away to   place   one   CL   that   TITLE.RESP Cham

       nơn   sua   laik...
       DET   VT   CMPZ
       that   to.ask that
       After running for a while, they arrive at a place, and an old Cham man asks
       them...

In addition to the above, a prepositional phrase (PP) or a noun phrase which consists of a
noun or two nouns has been used to modify a head noun. Apart from coming after the head
noun and before the determiner, it has not been possible to place them more accurately in the
noun phrase since there were no clauses in the data which included the prepositional phrase
or noun phrase along with an adjective, possessive noun phrase, or number–classifier
combination. In example (40) the NP is underlined and the PP that is modifying the noun
tanửh ea is in bold. Note that the NP in the PP—*play Ku*—consists of two nouns, with *Ku*
modifying the main noun *play*.

(40)   Nhu   hu   mayai laik   mơng   samăn dahlau   <u>tanửh</u>   <u>ea</u>
       PRO   TNS   VT   CMPZ   PREP   N   ADJ   N   <
       3   PAST   to.say that   from   time   in.the.past kingdom <

       <u>đi</u>   <u>play</u>   <u>Ku</u>   ni   sĕt   ta   tasik.
       PREP   **N**   **N**   DET   ADV   <   N
       in   **country** **Cambodian** this   entirely <   sea
       They say that formerly Cambodia here was entirely sea.

In example (41) a location NP consisting of two nouns—*chơk đangrêk* 'mountain
Dangrek'—modifies the head noun, *takai* 'foot'.

(41)   Maka   moh   labik tasik nơn di   <u>takai</u> **chơk**   <u>Đangrêk</u>   nơn
       CNJS   PREP   N   N   DET PREP   N   **N**   **N**   DET
       then   at   area   sea   that at   foot   **mountain** **Dangrek** that

       mada   ha   sa   koh,   rang   iau   Koh   Gôk   Dalôk.
       VN   NUM   CL   N   N   VT   N   N.PROP   N.PROP
       there.be   one   CL   island people to.call Island Gok   Dalok
       In the sea, near the foot of the Dangrek Mountains, there was an island; the
       Khmers called it Gok Dalok Island.

In a few examples, the head noun follows the number and classifier. This is apparently
found in older stories, in speech told to make them sound older, or in the speech of older
speakers of Cham. This word order, with the noun coming after its modifiers, is similar to
the word order of the noun phrases in Eastern (Vietnamese) Cham.

3.1.2 POSSESSION

Possession is shown by placing a noun phrase after the noun that is possessed. If an adjective exists in the main noun phrase, it may come before or after the possessive noun phrase (examples 42, 43). It is possible that the possessive noun phrase is restricted in form, although the limited amount of data examined for this sketch is not conclusive one way or the other.

(42) Nhu   bôh   _sang_ **hlŭn**   _prung_ klau   bôh.
      PRO   VT   N   **PRO**   ADJ   NUM   CL
      3   to.see house **1SG.LORESP**   big   three CL
      He saw my three big houses.

The possessive noun phrase and the adjective may also be reversed, with no apparent change in meaning.

(43) Nhu   bôh   _sang_ _prung_ **hlŭn**   klau   bôh.
      PRO   VT   N   ADJ   **PRO**   NUM   CL
      3   to.see house big   **1SG.RESP**   three CL
      He saw my three big houses.

3.2 VERB PHRASES

The verb phrase in Western Cham contains several elements. The head of the verb phrase is of course a verb. The verb may be intransitive, transitive, bitransitive, or non-active.

Three categories of words may precede the intransitive, transitive, and bitransitive verbs in the verb phrase: tense (TNS—hu 'past', hi 'future') (example 44), auxiliary (AUX—djauk 'must') (examples 45, 46) and a preverb adverb (PVA) (examples 45, 47).

(44) Nhu   **hu**   _mayai_ laik   mơng   samăn dahlau   tanŭh   ea
      PRO   **TNS**   VT   CMPZ   PREP   N   ADJ   N   <
      3   **PAST** to.say   that   from   time   in.the.past territory <

      di   play   Ku   ni   sĕt   ta   tasik.
      PREP   N   N   DET   ADV   <   N
      in   country   Cambodian this entirely <   sea
      They say that formerly the territory of Cambodia here was entirely sea.

(45) Yah   rôk   nao, bôh   drăp   kau,   hu   djauk   **ta**
      ADVZ   VT   DIR   VT   N   PRO   PRO   AUX   PVA
      if   to.dig away to.see thing 1.HIRESP 2.LORESP **must** **only**

      bray ni   ka   kau   wơk.
      VT   DET   PREP   PRO   FIN
      to.give this to   1.HIRESP back
      If you go dig it up and find my things, you have to give it back to me.

(46)   *Rani*   **djauk**   *hư*   *bray*   *pagŏl*   *tanưh*   *ea*   *ni*   *mai*
           ADV   **AUX**   PRO   VT   VT   N   <   DET   DIR
           now   **must**   2.LORESP   to.give   to.hand.over   territory   <   this   toward

           ka   *kau*   *wơk.*
           PREP   PRO   FIN
           to   1.HIRESP   back
           Now you must hand over this territory back to me.

(47)   *Miưk*   **dêl**   *ngăk*   *sang*   *ray?*
           N   **PVA**   VT   N   YN.QM
           younger.uncle   **ever**   to.make   house   also
           Have you (young uncle) ever built a house?

Two categories of words occur after the verb, directional (DIR—*nao* 'away', *mai* 'come') (examples 48, 50, 60, 65), and adverbs (ADV) (example 49, 50). Directionals generally occur immediately after the verb. Adverbs occur after the directionals.

(48)   *Yah*   <u>*rôk*</u>   <u>*nao,*</u>   *bôh*   *drăp*   *kau,*   *hư*   *djauk*   *ta*
           ADVZ   VT   **DIR**   VT   N   PRO   PRO   AUX   PVA
           if ·   to.dig   **away**   to.see   thing   1.HIRESP   2.LORESP   must   only

           bray   *ni*   *ka*   *kau*   *wơk.*
           VT   DET   PREP   PRO   FIN
           to.give   this   to   1.RESP   back
           If you go dig it up and find my things, you have to give it back to me.

(49)   *Hu*   *dok*   **sanăng**   **sanea**   *găn*   *lamư*   *rai*   *patao*
           TNS   VI   **ADV**   <   PREP   NUM   N   N
           PAST   to.live   **peacefully**   <   for.the.duration.of   five   reign   king

           păk   *Gôk*   *Dalôk*   *nơn.*
           PREP   N.PROP   N.PROP   DET
           at   Gok   Dalok   that
           They lived peacefully for five generations of kings at Gok Dalok.

(50)   *Nhu*   *đuaik*   **nao**   **tah**   *yơ.*
           PRO   VI   **DIR**   **ADV**   FIN
           3   to.run   **away**   **far**   already
           They have run far already.

The non-active verb appears alone (example 9), if it is there at all. Sometimes there is no verb in a non-active clause (examples 6–8).

3.3 ADJECTIVE PHRASES

Adjective phrases in WCham have an obligatory adjective (examples 51, 52) followed by an optional intensifier (usually *lô* 'very') (example 52).

(51) *Katiang ni   prung.*
      N   DET   ADJ
      boil   this **big**
      This boil is big.

(52) *Katiang ni   prung lô.*
      N   DET   ADJ   INT
      boil   this **big** **very**
      This boil is very big.

3.4 ADVERBIAL PHRASES

Phrases that are adverbial in nature have three different structures: adverb phrases, prepositional phrases, and a limited set of noun phrases.

3.4.1 ADVERB PHRASES

Adverb phrases are made up of an obligatory adverb as head (examples 53, 54, 67) and an optional intensifier (usually *lô* 'very', occasionally *đay* 'very') (example 54).

(53) *Ông   non   hamĭt   nhu   mayai bangi   păng.*
      N   DET   VT   PRO   VT   **ADV**   <
      TITLE.RESP that to.hear 3   to.talk **melodious** <
      He heard them speaking melodiously.

(54) *Madô laik mong dray mada   brŭk rawăm   lô   đay   dray*
      CNJS   <   PREP PRO   VT   N   ADJ   **ADV** INT   PRO
      but   <   from 1PL   there.be work to.be.busy **much** **very** 1PL

      ngăk   saphou ni   dhĭt   *ô.*
      VT   N   ADJ   **ADV**   FIN
      to.make   book   this   **completely.gone**   NEG
      Because we were so busy we weren't able to finish the book.

3.4.2 PREPOSITIONAL PHRASES

Prepositional phrases are adverbial in nature and can modify a verb or a noun. In example (55) the larger prepositional phrase, indicating where something was written, is in bold, the noun phrase that is part of the prepositional phrase is underlined, and the embedded prepositional phrase that modifies the noun *tapŭk* 'book' is double underlined.

(55) **Dalăm** <u>*tapŭk*</u> <u>**Ku**</u>   <u>*mong*</u> <u>*kal*</u> <u>*dahlau*</u> *nhu* *hu*
      PREP   N   N   PREP   N   ADJ   PRO   TNS
      in   book   Cambodian   **from**   time   **in.the.past** 3   PAST

      hrăk,   *ngĕn* *saphou* *Preah*   *Thong*   *Neang*   *Neak.*
      VT   N   N   N.PROP   N.PROP   N.PROP   N.PROP
      to.write   name   book   Preah   Thong   Neang   Neak
      This is written in an ancient Khmer book called 'Preah Thong Neang Neak'.

3.4.3 NOUN PHRASES

There is a limited set of noun phrases that can be used adverbially. The head noun of these noun phrases often relates to time or location, such as *malam* 'night', *mabroi* 'yesterday', and other time-related nouns, as well as a few words like *sang* 'house, home' (examples 56–58).

(56) | *Tăl* | *păng* | *nao* | **malăm** | **di** | **sa,** | *dăm* | *nơn* | *hamĭt* | *ông* | *hŏng* |
|------|--------|-------|-----------|--------|---------|-------|-------|---------|-------|--------|
| ADVZ | VT | VI | N | **ADJ** | < | N | DET | VT | N | PREP |
| when | to.listen | to.go | **night** | **first** | < | suitor | that | to.hear | old.man | with |

muk	*mayai*	*gauk*	*laik:*	**"Pagê**	**ni**	*ông*	*khĭn*
N	VT	REFL	CMPZ	N	**DET**	N	VT
grandmother	to.speak	together	that	**tomorrow**	**this**	old.man	to.want

padăr	*dăm*	*nơn*	*ngăk*	*jiang*	*jal*	*meok*	*nao*
VT	N	DET	VT	VN	N	VT	VI
to.do.again	suitor	that	to.do	to.be	large.fish.net	to.carry	to.go

trah	*kan*	*dalăm*	*bơng*	*ni*	*bơng*	*nơn".*
VI	F	PREP	N	DET	N	DET
to.fish.with.net	fish	in	lake	this	lake	that

While he was listening the first night, the suitor heard the man and woman talk together, saying: "Tomorrow I want to use that suitor as a fish net which I will take to fish in such and such a lake".

(57)	*Marekăng*	*mai*	**mabroi.**
N	VI	N
American(s)	to.come	**yesterday**

The Americans came yesterday.

(58)	*Hlŭn*	*hi*	*nao*	**sang.**
PRO	TNS	VI	N
1SG.LORESP	FUT	to.go	**house**

I will go home.

4. SENTENCES

4.1 QUESTIONS

4.1.1 YES/NO QUESTIONS

Yes/no questions in Cham can be formed in at least three ways, all of which involve adding a word or a phrase to the end of the clause.

The first way is by adding the word *ray* to the end of the clause (example 59). In other locations in the sentence, *ray* can mean 'or', 'so', or 'also'. In this type of yes/no clause, *ray* marks the sentence as a question. In addition, the intonation, which usually falls at the end of a sentence, rises at the end of yes/no questions.

(59)	*Bôh*	**ray?**
VT	FIN
to.see	YN.QM

See?

The second type of yes/no question adds *mǐn* to the end of the clause (example 60). It can also be used at the end of a statement as an affirmative particle, but in the following clause, it marks the clause as a yes/no question. Rising intonation at the end of the clause indicates this is a question.

(60)

Chǐm	*kung*	*rang*	*chuh*	*nɔn*	*rang*	*mɔk*	*mai*
N	REL.PRO	PRO	VT	DET	PRO	VT	DIR
animal	that	3.INDEF	to.hunt	that	3.INDEF	to.bring	toward

bɔng	*hu*	*mǐn?*
VT	VT	**FIN**
to.eat	to.be.able.to	**YN.QM**

Can they eat animals that they hunted?

The third type of yes/no questions uses a phrase at the end: *ray ha soh mǐn*. This seems to function as a tag ending on the sentence, as in English 'Bill is tall, isn't he?' For example:

(61)

Rang	*dêl*	*padǎr*	*sau*	*nao*	*doah*	*pagui*	*chǐm*	**ray**
PRO	ADV	VT	N	VI	VT	VT	N	**ADV**
3.INDEF	ever	to.use	dog	to.go	to.search.for	to.chase	animal	**also**

ha	*soh*	*mǐn?*
CONJ	**FIN**	<
or	**or.not**	<

Do they or don't they ever use dogs to search for, to chase the animals?

4.1.2 CONTENT QUESTIONS

4.1.2.1 WH QUESTIONS

Content questions are questions that ask for more than a yes or no answer. Typically, these questions use a word that replaces part or all of the noun phrase, adverb phrase, adverbial clause and so on. In some cases, the questioned part of the sentence is moved to the front of the sentence.

In the following non-active sentence, the type of the subject noun is questioned (i.e. the quality, type, etc.). The noun *chǐm* is followed by the content question word, *gêk*.

(62)

Chǐm	*gêk*	*manǔng*	*kung*	*rang*	*nao*
N	**WH.QM**	ADJ	REL.PRO	PRO	VI
birds/animals(generic)	**what**	some	that	3.INDEF	to.go

chuh	*nɔn?*
VT	DET
to.hunt	that

What are the birds (animals) that they go hunting?

In example (63) the direct object is questioned. The word *hagêk* replaces the direct object in the sentence following the verb.

(63)

Rang	*padǎr*	**hagêk**	*manǔng*	*samrap*	*chuh*	*chǐm*	*nɔn?*
PRO	VT	**WH.QM**	ADJ	CONJ	VT	N	DET
3.INDEF	to.use	**what?**	some	for.the.purpose.of	to.hunt	animals	that

What (material) do they use to hunt animals?

In example (64) the manner is questioned. The content question word (or actually phrase) is *yau băr* and it occurs in the position in the sentence where adverbs or adverbial clauses explaining manner would be found.

(64)   *Kanrãm*   *ko   rang   chĕk   yau   băr?*
       N   FOC   PRO   VT   **WH.QM** <
       animal.trap(falls on victim) FOC   3.INDEF to.place **how?**   <
       How do they set up the trap that falls on victims?

4.1.2.2 EITHER/OR

Either/Or questions are content questions that give the one who responds only two options (examples 65, 66). In this type of question, there are two clauses which are connected by the conjunction *ha* 'or'. The conjunction is marked by bold type, and the full clauses are underlined on either side of the conjunction.

(65)   *Rang   nao   chuh   nŏn   rang   nao   chuh   ma-ĭn   ha*
       PRO   VI   VT   DET   PRO   VI   VT   ADV   **CONJ**
       3.INDEF to.go   to.hunt that   3.INDEF to.go   to.hunt for.pleasure **or**

       rang   chuh   mŏk   mai   bŏng?
       PRO   VT   VT   DIR   VT
       3.INDEF to.hunt to.bring toward to.eat
       When they go hunting, do they hunt the animals for pleasure or do they hunt them for food?

(66)   *Hi   nao   ha   rang   doa   rang   ha   nao   lô   gauk?*
       TNS   VI   NUM   CL   NUM   CL   **CONJ** VI   ADJ   N
       FUT   to.go   one   person   two   person   **or** to.go many   companion
       Do they go alone, in a group of two, or a lot more (than that)?

4.2 COMMANDS

There are two types of commands that have been found so far. The first is a mild command or request. It begins with *lakau* 'to ask' and ends with *wek* 'mild imperative'. For example:

(67)   **Lakau** *chŏp   đŏ   ni   wek!*
       VT   VT   ADV   <   **FIN**
       to.ask to.stop at.this.point < **MILD.IMPER**
       Please let me stop now! (lit. I ask to stop now.)

The second is a negative command. In this type of command, the word *di* is found between the noun phrase subject and the verb phrase and *juai* 'negative imperative' is found at the end of the clause (examples 68–70).

(68)   *Hŭ   di   tăk   juai.*
       PRO   **NEG...** VT   **FIN**
       2.RESP **don't** to.cut **NEG.IMPER**
       Don't cut (me) down.

The subject of the negative command is optional. For example:

(69)   **Di**   *đuaik* *juai.*
      NEG... VI   **FIN**
      don't   to.run **NEG.IMPER**
      Don't run away.

A regular command may be given in which the main clause (underlined) ends with the final particle *wek*. For example:

(70)   *Hơi,*   *rup* *hư*   *di*   *nao*   *hlay* *juai,*   <u>*hư*</u>
      INTER N   PRO   NEG... VI   PRO   FIN   PRO
      Oh   body 2.LORESP don't to.go where neg.imper   2.LORESP

      <u>*chang*</u>   <u>*kau*</u>   <u>*moh*</u> <u>*ni*</u> <u>**wek**</u>.
      VT   PRO   PRO   <   **FIN**
      to.wait.for 1.HIRESP here   <   **MILD.IMPER**
      Hey, body, don't go anywhere, just wait for me here.

APPENDIX: LIST OF ABBREVIATIONS

<	Connects to previous word to form compound word
ADJ	adjective
ADV	adverb
ADVZ	adverbialiser
AUX	auxiliary
CL	classifier
COMP.ADV	comparative adverb
CMPZ	complementiser
CONJ	conjunction
CNJS	conjunction, sentence level
DET	determiner
DIR	directional
DUM.SUBJ	dummy subject
FIN	final particle
FOC	focus marker
HIRESP	higher respect
IMPER	imperative
INT	intensifier
INTER	interjection
LORESP	lower respect
N	noun
N.PROP	noun, proper
NEG	negative
NUM	number
PART	particle
PREP	preposition

PRO	pronoun
PVA	preverbal adverb
REFL	reflexive
REL.PRO	relative pronoun
TNS	tense
U	unknown
VB	verb, bitransitive
VI	verb, intransitive
VN	verb, non-active
VT	verb, transitive
WH.QM	WH question marker
YN.QM	yes/no question marker

CHAM EVIDENCE FOR KHMER SOUND CHANGES

ROBERT K. HEADLEY, JR

1. INTRODUCTION

When a recipient language borrows, for whatever reason, a word from a donor language, that word in the recipient language ceases to undergo any historical changes which take place in the donor language. It becomes a kind of fossil and may tell us much about the phonology of the donor language at the time the borrowing took place. Consider the Modern German reflexes of Latin /k/ in the words *Kiste* 'chest' (from Latin *cista*) and *Zelle* 'cell' (from Latin *cella*). The word *Kiste* must have been borrowed earlier, before Latin /k/ underwent palatalisation before front vowels, and *Zelle* must have been borrowed later, after some degree of palatalisation had taken place giving German /ts/. Proto Mon-Khmer seems to have had at least two series of initial stops, one voiceless and the other voiced. Huffman (1976) suggests that there were five possible stages in the development of Mon-Khmer phonetic systems. These are: (1) voiceless: voiced distinction in initial stops; two-way distinction in initial continuants; as yet little or no vowel differentiation; (2) simultaneous change in the articulation of one set of initial consonants and development of allophonic variation in following vowels, still in complementary distribution vis-a-vis two distinctive sets of initials; (3) register becomes phonemic due to a complete merger at some point in the system, probably in the continuants, with retention of allophonic distinction in initials in complementary distribution vis-a-vis first and second register vowels; (4) complete merger of initial stops, with complete register dichotomy in the vowels; (5) loss of register contrast due to change in vowel position and diphthongisation. It is generally agreed that Khmer inherited a series of voiced stops from Proto Mon-Khmer /*a, *d, *j, *g/. Khmer retained these voiced stops until some time after AD 1500, when they became devoiced. It is difficult to put a precise date to any sound-change since changes are very gradual. Jenner (1976:694) believed that "This [devoicing] process appears to have reached its culmination between the sixteenth and the eighteenth centuries". Lewitz (1967:387 ff.) believed that the devoicing took place between the sixteenth and seventeenth centuries and was complete by the eighteenth century. Lewitz based her argument on the way the name Angkor (Modern /ʔaŋkɔɔ/, Middle Khmer /ʔɔŋgɔɔr/) was written by foreigners. Sixteenth century Spanish and Portuguese travellers wrote *Angar, Angor, Angon*, apparently indicating that the *g* was still voiced. On the other hand, seventeenth century visitors to Cambodia wrote it *Anckoor* and *Onco*. Vickery (1992:244–247) disagreed with Lewitz. He believed, based on his study of the Portuguese and Spanish renderings of Khmer words as recorded in Groslier (1958), that there was ample evidence from the Portuguese and Spanish records that devoicing had occurred by the last two decades of the sixteenth century. There was also some major restructuring of the Khmer vowel system at some point, probably after the devoicing of the

---

initial stops. An examination of the phonology of loan words from Khmer into Western Cham can illuminate some of the phonology of earlier stages of Khmer before the initial stops were devoiced.

2. HISTORICAL BACKGROUND

The Chams have been in contact with Khmer-speaking communities for at least 14 centuries. During this time there were numerous economic, military, and political contacts; certainly the languages interacted. The Chams were historically located in several enclaves[1] along the coast of present-day Vietnam between Cape Mui Ron in the north and Bien Hoa in the south, and the Khmers were located to their south and west. Briggs (1951:13) states that "The Chams seem to have held the Mekong from the mouth of the Khong at Stung Treng up to near the mouth of the Mun River until the coming of the Khmers". He believes that the Khmers either drove the Chams out of the area around the Mun River mouth or absorbed them. The Chams were under constant political and military pressure from first the Chinese and later the Vietnamese in the north, and from the Khmers in the south and west. Sometime after 1471, when Vijaya (Binh Dinh), the last major Cham enclave, fell to the Vietnamese, some of the Chams fled west into lands controlled by the Khmers and settled in small villages, often along the banks of major rivers, scattered throughout the country.[2] Today the descendants of these Chams are known as the Western Cham. A second group, the Eastern Cham, remained in their homeland around the cities of Phan Rang and Phan Thiet. At some point, probably by at least the tenth century AD (Maddieson and Pang 1993:75), a third group of Cham took to the sea and went north to Hainan Island, where they now form a minority known as Utsat. The languages of each of these groups have developed in different directions. Western Cham has become a register language. Utsat has become a tonal language and Eastern Cham is developing a tonal system. All of these phonetic changes in the three Cham dialects seem to have developed as a result of the devoicing of initial consonants.

The lack of indisputable borrowings from Khmer in Eastern Cham and Utsat suggests that there were no extensive borrowings from Khmer prior to the break-up of Champa despite the long period of contact. The situation with Western Cham is different; the current language is filled with borrowings from Khmer. The borrowed words are not limited to nouns and verbs, but also include some verbal and nominal auxiliaries. These borrowings appear to have been made over a long period of time and may be conveniently divided into two sets. The first set contains old borrowings made before Khmer devoiced its voiced bilabial, dental, palatal, and velar stops. The second set contains later borrowings after the devoicing. The first set is the subject of this paper.

The topic of loan words in Southeast Asian languages is a thorny one, and often it is difficult to determine the direction of the borrowing. For this study, forms from three dialects

---

[1]   Higham (1989:297 ff.) and Taylor (1992:155) doubt the long-term existence of a single Cham political entity, noting that the Chams themselves recognised several polities centred on river-mouth urban centres. These included Indrapura, centred around Tra-Kieu, Vijaya, centred around Binh-Dinh, and Kauthāra, centred around Nha-Trang.

[2]   Major Cham settlements in Cambodia were located around Battambang City, on the southern side of the Tonle Sap; around Kampot City, along the Tonle Sap River above and below Phnom Penh; and along the Mekong River above and below Kompong Cham City.

of Cham were compared.[3] Obvious Indic loan words were discarded, as were obvious Mon-Khmer borrowings that appeared in all three dialects. If a word that seemed to be a borrowing from Khmer occurred in KGT Cham or WCD Cham but not in Eastern Cham, it was assumed that it was a borrowing that had taken place after the break-up of Champa, or roughly after AD 1500.

3. PHONOLOGY OF KHMER AND WESTERN CHAM

The phonologies of Khmer and Cham are very similar; they share many of the same sounds, as a comparison of the systems will show.

Three historical periods are recognised for Khmer.[4] Old Khmer (OK) from about AD 600 to 1431, Middle Khmer (MK) from about AD 1431 to 1700, and Modern Khmer from about 1700 on. Old and Middle Khmer had the following phonemic inventory[5]:

TABLE 1: OLD/MIDDLE KHMER

CONSONANTS:

Voiceless stops:	p		t		c	k	?
Voiced stops:	b	?b	d	?d	j	g	
Nasals:	m		n		ñ	ŋ	
Voiced liquids:			r	l			
Voiceless spirants:			s				h
Semivowels:	w		y				

VOWELS:

i	ii	ɨ	ɨɨ		u	uu
e	ee	ə	əə		o	oo
ε	εε					
		a	aa		ɔ	ɔɔ

DIPHTHONGS:

iə	iiə
ɨə	ɨɨə
uə	uuə

After considerable restructuring, Modern Standard Khmer has emerged with the following inventory of phonemes:

---

3 These are the Western Cham dialects of Kompong Thom Cham (KGT), spoken in southern Kompong Thom Province about 50 miles north of Phnom Penh; Chau Doc Cham (CHD), spoken around Chau Doc in Vietnam and in neighbouring Cambodia south of Phnom Penh; and Eastern Cham, spoken between Phan Rang and Phan Thiet in south central Vietnam.

4 Diffloth (1992:271–272); for a somewhat different dating of the three periods, see Jenner (1969:3), which is based on Jacob. Ferlus (1992) divides the language of the Khmer inscriptions into Pre-Angkorian Old Khmer seventh and eighth centuries, Angkorian Old Khmer ninth–thirteenth centuries, and Middle Khmer fourteenth–eighteenth centuries.

5 Based on Jenner and Pou (1980–81:xviii–xix).

TABLE 2: MODERN STANDARD KHMER

p	t	c	k	ʔ
b	d			
(f)[6]	s			h
v	(z)	(ž)		
		y		
m	n	ñ	ŋ	
	r			
	l			

i	ii	ɨ	ɨɨ	u	uu
e	ei [eˆː]	ə	əə	o	ou [oˆː]
	ee [eˇː]				oo [oˇː]
	εε			ɔ	ɔɔ
a	aa			a	aa

ie	ae	ao	aə	ea	ɨə	uə	oa	wə

In some dialects there is a suprasegmental phoneme of register that is manifest by tense, clear, open, higher pitch vowels in syllables following originally voiceless consonants (= high register) and lax, breathy, close, lower pitch vowels in syllables following originally voiced consonants (= low register). Some of the vowel changes that took place after the devoicing of originally voiced stops are given below:

TABLE 3: KHMER VOWEL CHANGES

Original vowel nucleus	High register	Low register
a'	u/wə	a
a	ɔɔ	aa
ā'	ea/oa	a
ā	ie	aa
i	ɨ	ə
ī	ii	əy
ɨ	ɨ	ə
ɨ̄	ɨɨ	əɨ
u	u	o
ū	uu	oo
œ	əə	aə
e	ei	ee
ē	εε	ae
o	ou	ao

Old/Middle Cham had a system nearly identical to Old/Middle Khmer with the probable addition of /ʔj/, a preglottalised /j/. A provisional list of Old Cham phonemes is given below.

---

6   The phonemes in parentheses occur only in borrowed words.

TABLE 4: OLD/MIDDLE CHAM

CONSONANTS:

Voiceless stops:	p		t		c		k	(*?*)
Voiced stops:	b	*?b*	d	*?d*	*j*	*?j*	g	
Nasals:	m		n		ñ		ŋ	
Voiced liquids:			r	*l*				
Voiceless spirants:			s				h	
Semivowels:	w		y					

VOWELS:

i	*ii*				*u*	*uu*
e	*ee*	*ə*	*əə*		*o*	*oo*
ε	*εε*					
		a	*aa*		*ɔ*	*ɔɔ*

DIPHTHONGS:

iə	*iiə*
uə	*uuə*
ao	

Both phonetic systems, Old/Middle Khmer and Old Cham, underwent some radical changes on their way to the modern languages. Among many other changes, both have devoiced their voiced stops. In Khmer this devoicing resulted in the formation of some new vowel nuclei as presented above. In Western Cham it resulted in the development of a register system. High register (or second register) vowel nuclei in modern Western Cham are those which follow originally voiced stops; they generally have a lower pitch than those which follow originally voiceless stops, are higher in tongue height, and have a 'breathy' quality. Low register (or first register) vowel nuclei are those which follow originally voiceless stops; they have a higher pitch, are lower in tongue height, and sound 'clearer' or 'sharper' than the nuclei following originally voiced stops; they may also have a lower onglide.

TABLE 5: MODERN WESTERN CHAM

CONSONANTS:

Voiceless stops:	p	t	c	k	?
Voiced stops:	*?b*	*?d*	*?j*		
Nasals:	m	n	ñ	ŋ	
Voiced liquids:		*l*		r	
Voiceless spirants:		s		h	
Semivowels:	w		y		

VOWELS:

i	*ii*	*ɨ*	*ɨɨ*	*u*	*uu*
e					
		ə	*əə*	*o*	*oo*
ε	*εε*				
		a	*aa*	*ɔ*	*ɔɔ*

Additional to these are the glides /ia, ii, ea, au, ua, oa/ and a suprasegmental phoneme of register. The first (or low) register is unmarked while the second (or high) register is indicated in this paper by underlining.

4. WESTERN CHAM BORROWINGS

The dialects of Western Cham contain a significant number of apparent loan words from Khmer. Many of these words, for example CHD Cham *gre* 'bed' and KGT Cham *smôk ~ grôk* 'be dirty' are probably recent borrowings. Other words are borrowings from Indic sources which may or may not have been through Khmer as an intermediary, for example CHD and KGT Cham *bhik* 'Buddhist monk'.[7] A third group of words may be very ancient borrowings from some Austroasiatic source or may be from a common protoform, if indeed there is a genetic relationship between Austroasiatic and Austronesian. Examples of words in this category include KGT Cham *bhay* 'otter', KGT and CHD Cham *pabe* 'goat', and KGT and CHD Cham *gǎr* 'drum'. A fourth group of borrowings are those that are very early borrowings from Khmer made before the devoicing of the Khmer initial stops. This group will be discussed below.

Khmer *b* = Cham *b*

    buta /put/ 'falseness' KGT *bǔt* [pu̯t][8]

    trabamnga /trɑpeaŋ/ 'pond' KGT *tabǎng* [tapa̯ŋ]

    lbā'ka /lpeak/ 'kind of rattan' KGT *labǎc* [lapa̯k]

    brama /prɔɔm/ 'agree' KGT *brom* [prɔ̯:m]

Khmer *d* = Cham *d*

    dā'ta /toat/ 'kick' KGT *dět* [tɛt] [9]

    dā'la /toal/ 'be stumped' KGT *děl* [tɛl]

    dā'na /toaŋ/ 'on time' KGT *děn* [tɛn]

    dā'ka /teak/ 'trap' KGT and CHD *dǎk* [ta̯k]

---

7   Other words of this type are CHD *bual* 'army', KGT *bêl* 'time', KGT and CHD *bet* 'medicine', KGT *duih* 'fault', and KGT and CHD *dop* 'army'.

8   Khmer forms are given first in a transliteration used by Jenner and Pou (1980–81) and based on the traditional Anglo-American system for romanising the devanagari writing system and then in a modern phonemic version enclosed in slants; the Western Cham forms are given first in a phonemic form followed by a phonetic form in square brackets.

9   Old Khmer *$\bar{a}$' became /ɛ/ before final dentals in Western Cham. Since the phoneme /ɛ/ is rare in Modern Cham, could this mean that there was already some allophonic change in the Khmer vowel *$\bar{a}$' /C _____ C   when the borrowing took place?

    [+voice]   [+cor]

    [-son]   [+ant]

                [-cont]

dāhāna /tiehien/ 'soldier' KGT and CHD dahan [tₐhₐːn] [10]

dram /troam/ 'endure' KGT drŏm [trɔm]

drunga /truŋ/ 'cage' KGT and CHD drŭng [truŋ]

Khmer *j = Cham *j

jāla /ciel/ 'kind of basket' KGT and CHD jal [jaːl] 'cast net'

jamnwsa /cumnuəh/ 'substitute' KGT jamnuaih [jₐmₙuaih] or [jumnuaih]

jhnah /cneah/ 'win' KGT januɾh [canih] CHD chanuɾh [canih] [11]

pañjūna /bañcuːn/ 'send' KGT pajun [pacuːn]

jam /coam/ 'bruised' KGT jŏm [cɔm]

jhlī /cliː/ 'rub' KGT jali [jₐliː]

jrula /crul/ 'exceed' KGT jrŭl [crul]

jra'ka /crwək/ 'pickle' KGT jrŏc [crɔk]

Khmer *g = Cham *g

ga'la /kwəl/ 'stump' KGT and CHD gŏl [kɔl] 'trunk/stump'

gara /kɔː/ 'kapok' CHD go [kɔː]

guka /kuk/ 'jail' KGT and CHD gŭc [kuk]

gā'ta /koat/ 'he, she' KGT and CHD gĕt [kɛt]

greca /krɨc/ 'sprain' KGT grĕik [kₐɾɛik] or garĕk [kₐɾɛk]

grā'na /kroan/ 'enough' KGT gĕn [kɛn]

gum /kum/ 'grudge' KGT gŭm [kum]

Crucial to the argument put forth in this paper is the assumption that, when a language borrows a word, some attempt will be made to reproduce the pronunciation of the donor language as accurately as possible. In this way the pronunciation of the word at the time it was borrowed will be preserved or fossilised in the borrowing language. For example, if Western Cham borrowed the word for 'field' as val /vaːl/ from Khmer, it was probably, at the time of borrowing, pronounced very much like /vaːl/ by the Khmers. The evidence above suggests that Western Cham borrowed a large number of words from Khmer before the devoicing of the initial stops. It also suggests that the restructuring of the Khmer vowel

---

10   Jenner and Pou (1980–81:389) derive Khmer (KH) /tiehien/ from the MK base hānā. Eastern Cham apparently does not have this word, but it occurs in Rade and Jarai as than and tɤhan respectively.

11   The /ɨ/ in Western Cham is usually a reflex of Proto Chamic (PC) *ā ; in this word it may be due to the preceding nasal.

system took place after the devoicing. If the diphthongisation of the Khmer *ā* had taken place before the borrowing and therefore before the devoicing, we would expect the Western Cham form to be something like **veal*. There is additional evidence from other Western Cham borrowings from Khmer that the borrowing was made before the restructuring of the Khmer vowel system, for example *tanot* /tano:t/ 'sugar palm' (KH *tnota* /tnaot/), *kamot* /kamo:t/ 'ghost' (KH *khmoca* /kmaoc/), *ran* /ra:n/ 'platform' (KH *rāna* /rien/), *tơ* /tə:/ 'only' (KH *te* /tae/), *kĭn* /kin/ 'to mill rice' (KH *kina* /kən/), *yam* /ya:m/ 'to patrol' (KH *yāma* /yiem/). In later borrowings from Khmer by Western Cham, the vowel modifications which took place in Khmer after the devoicing of the initial stops are reflected, for example *jeang* /cɛaŋ/ 'more than' (KH *jānga* /cieŋ/), *bêl wêlea* /pe:l ve:lea/ 'time' (ultimately from an Indic source through KH *bēla vēlā* /pe:l ve:lie/).[12]

5. CONCLUSIONS

Based on the study of a corpus of Western Cham borrowings from Middle Khmer, it appears that at the time these borrowings were made Middle Khmer still had voiced initial stops. This stage probably equates with Huffman's Stage 1. Western Cham borrowed these words with the voiced stops intact. These then joined native initial voiced stops in their regular historic development in Cham. That the Khmer vowel nuclei following the original voiced stops were maintained in the Western Cham borrowings with little or no change suggests that the restructuring of the Khmer vowel system had not yet begun. It is very difficult to determine the date these borrowings were made, but, based on their absence in Eastern Cham, it is believed that the bulk of the borrowings took place after the break-up of Champa in the mid–late fifteenth century AD and the separation of Western and Eastern Cham. If this is an accurate date, then Standard Khmer passed through Huffman's five stages in less than 400 years.

REFERENCES

Briggs, Lawrence Palmer, 1951, *The ancient Khmer empire.* Philadelphia: The American Philosophical Society.

Diffloth, Gérard, 1992, Khmer. In William Bright, ed. *International encyclopedia of linguistics*, 271–275. New York and Oxford: Oxford University Press.

Ferlus, Michel, 1992, Essai de phonétique historique du khmer. MKS 21:57–89.

Groslier, Bernard Philippe, 1958, *Angkor et le Cambodge au XVIe siècle d'après les sources portugaises et espagnoles.* Paris: Presses Universitaires de France.

Higham, Charles, 1989, *The archaeology of mainland Southeast Asia.* Cambridge: Cambridge University Press.

---

[12]   The diphthong /ea/ in recent Western Cham borrowings is a reflex of original Khmer *ā* following a voiced consonant and is clearly distinct from the diphthong /iə/, which reflects original *ia*; for example *phian* [phiən] 'opium' (ultimately from Chinese through KH *phiana* /phiə n/), *chakial* [cəkiəl] 'to scrape' (KH *chkiala* /ckiəl/).

Huffman, Franklin E., 1976, The register problem in fifteen Mon-Khmer languages. In Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, eds *Austroasiatic studies*, 575–589. Honolulu: The University Press of Hawaii.

Jenner, Philip N., 1969, *Affixation in modern Khmer*. Doctoral dissertation, University of Hawaii.

1976, The relative dating of some Khmer Cpā'pa*. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta, eds *Austroasiatic studies*, 693–710. Honolulu: The University Press of Hawaii.

Jenner, Philip N. and Saveros Pou, 1980–81, *A lexicon of Khmer morphology*. MKS 9–10. Honolulu: The University Press of Hawaii.

Lewitz, Saveros, 1967, La toponymie khmère. *Bulletin de L'Ecole Française d'Extreme Orient* 53/2:375–451.

Maddieson, Ian and Keng-Fong Pang, 1993, Tone in Utsat. In Jerold A. Edmondson and Kenneth J. Gregerson, eds *Tonality in Austronesian languages*, 75–89. Honolulu: The University of Hawaii Press.

Taylor, Keith W., 1992, The early kingdoms. In Nicholas Tarling, ed. *The Cambridge history of Southeast Asia*, vol.1, 137–182. Cambridge: Cambridge University Press.

Vickery, Michael, 1992, Loan words and devoicing in Khmer. MKS 28–29:240–250.

THE CONTRIBUTION OF CAT GIA ROGLAI TO CHAMIC

ERNEST W. LEE

1. INTRODUCTION

Northern Roglai (NR) *buŋā* and Cat Gia Roglai (CGR) *puwiŋ* both meaning 'flower' and NR *srā:ʔ* and CGR *tʰuŋʔ* both meaning 'crossbow' might not be considered as cognates for doing lexicostatistic studies and would certainly not suggest that the two Roglai languages are closely related. The word Roglai itself probably comes from Proto Chamic (PC) **ʔura:ŋ* 'person' + **dlai* 'forest'. In NR **ʔura:ŋ* is reduced to *ra:k* 'classifier for people' and to the prefix *ra-* [aa] 'they, he, she' which is widely used in narrative discourse where the referent is clear from the context. The Northern Roglai people refer to themselves as Radlai [aadlai]. The Cat Gia Roglai people also refer to themselves as Radlai [radlai]. Given the origin of the word Roglai, it would be possible for the term to refer to any language where the people were considered to be people of the forest. The words *buŋā* and *puwiŋ* as well as *srā:ʔ* and *tʰuŋʔ* are definitely cognate and NR and CGR are closely related languages. The purpose of this paper is to show some of the contribution that Cat Gia Roglai can make to the understanding of the Chamic languages.

Sections 2–8 focus on various phonological developments in CGR and their significance to the wider Chamic picture. Special attention is given to nasalisation in §4 since there is more to say about it and it potentially has significance for Acehnese as well as the mainland Chamic languages of Southeast Asia. In §10, I show how CGR *puwiŋ* developed from Proto Chamic (PC) **buŋa* and how *tʰuŋʔ* developed from PC **srā:p* (Lee 1966, Burnham 1976). Finally, in §11, I summarise the various CGR features shared with specific Chamic languages.

The CGR data on which this study is made is very sketchy and some of it not completely reliable phonetically. It comes from about two hours of contact with a CGR village in 1974 working under the auspices of the Summer Institute of Linguistics. Without my background in Northern Roglai, it would not have been possible for me to get even that amount of useful data in the short time available.

1.1 DEMOGRAPHIC BACKGROUND

Cat Gia Roglai is not a large language group. To the best of my knowledge in 1974 it was spoken in only two hamlets in Ninh Thuan province in Vietnam. The information given to me at that time was that there were 1,187 people in Ap [hamlet] Cat Gia and 888 in Ap Binh Nghia-Cham, both located in Xa [village] Cat Hai north of Phan Rang. I have arbitrarily

---

referred to the language as Cat Gia Roglai because the information I got was from that particular hamlet. The people of both hamlets were in regular contact with Vietnamese speakers and the Cham-speaking people of Ninh Thuan province. It is not certain how much contact they may have had with some of the other Roglai-speaking people, but was apparently less than with Cham speakers. It is unlikely that they would have had any significant recent contact with Northern Roglai. The dress of the CGR people was more similar to that of the Cham than to that of the majority of the Montagnard people, but this was also true of some of the people traditionally called Southern Roglai. I have not included CGR as a Southern Roglai dialect since it appears to be somewhat distinct as well as geographically isolated.

1.2 CAT GIA ROGLAI AS A ROGLAI LANGUAGE

As noted above, the CGR people refer to themselves as Radlai. Certainly in their minds and in the minds of the neighbouring Chams and Vietnamese they are Roglai.

CGR shares a number of features with Cham so that in many respects it looks more like Cham than Northern Roglai. I am less familiar with Southern Roglai, but superficially CGR appears to share more with Cham than any of the other Roglai dialects do, even though probably most Southern Roglai dialects are in regular contact with Cham. Contact between Northern Roglai and Cham, on the other hand, is virtually nil.

One problem with the data is frequent variation between what was probably good CGR and what is quite evidently Cham. Because of their regular contact with the more prestigious Cham and the informants being bilingual in Cham, they very likely did intersperse some Cham pronunciations and lexical items. A second problem is that I undoubtedly missed or misrepresented some things because I was listening and writing through a NR grid. One particular variation (inconsistency in hearing) was in recording reflexes of voiced stops which I noted variously as voiced, voiced aspirated, voiceless, and voiceless aspirated (see §2.1). In spite of the these problems, however, there is sufficient consistency in the data to be able to make a number of valid observations about CGR.

One phonological innovation which CGR shares with all other dialects of Roglai, but not with Cham or any of the other Chamic languages is the denasalisation of final nasal consonants to oral stops (except where the preceding vowel is nasalised) as: CGR *lapat* 'eight' and *tʰalapat* 'nine', with the *[]lapat* from PC **lapan* which is reflected variously in Chamic as a component of 'eight' and 'nine'; *brəp* from PC **brOm* 'arrow'; *ʔbək* 'eat' from PC **ʔbəh*. The CGR examples all represent what has happened following short protovowels. The same was undoubtedly true following long vowels where the final stops have subsequently been lost in CGR (see §3.1).

NR and CGR appear to share some features not shared by Cham or the other Roglai dialects. I have already mentioned that both groups refer to themselves as Radlai with the *dl* cluster and an *a* in the pretonic syllable. The term Roglai has most likely come into Vietnamese and other languages from Cham, where it would be pronounced [rəglai]. To my knowledge, of the Lowland Chamic languages, only NR and CGR have the cluster [dl]. The others all have [gl]. NR and CGR also share a common form *kamĩn* 'we (exclusive)' which reflects the final nasal vowel **ĩ* with nasal accretion as described in §4.1.3.

Lexically CGR shares the word for 'cooked rice' with NR. NR has *bu* and CGR has *p^hu* alternating with *bu* in my data. The other mainland Chamic languages use reflexes of PC *l̩_sɔi* for 'cooked rice' whereas reflexes of *bu:r* are used for 'rice soup', although from my data I am not certain of the distinction between Haroi *lasôi* and *p^hu* since I have both listed as meaning 'cooked rice'. (Acehnese also has *bu* for cooked rice.)

With further study, details of comparison of CGR with NR and Southern Roglai could be amplified, but the denasalisation of the final nasal consonants is probably the most significant identifying feature of all that is called Roglai. Otherwise, as already noted, there are many features of CGR which are more like Cham than NR and perhaps Southern Roglai dialects as well. See §11.

2. DEVELOPMENT OF WORDS WITH SYLLABLE-INITIAL VOICED STOPS

Initial voiced stops have played an extremely important role in the development of the Chamic languages. Thurgood (1995) pulls together the various developments in Chamic and discusses their significance.

CGR, very much unlike NR, seems to be moving in the same direction as the other Chamic languages where register and/or tone have developed from syllables and words beginning with original voiced stops. I treat this subject first since it involves such a large number of words and affects varying aspects of the discussion throughout the rest of this paper, especially with respect to how the syllable-initial stops as transcribed are to be interpreted.

2.1 THE CGR EVIDENCE

As already noted, the reflexes of PC voiced stops in CGR are variously transcribed in my data as voiced, voiced aspirated, voiceless, and voiceless aspirated. The reason for this is probably also varied. I have already mentioned that I was listening through a NR grid. Sometimes I may have written a voiced stop simply because that was what I was expecting, but clearly that does not account for all the variations, particularly when NR as a matter of fact does have all four possibilities, but contrastively. That is, NR contrasts the sets: *p*, *ph*, *b*, *bh*; *t*, *th*, *d*, *dh*; *c*, *ch*, *j*, *jh*; and *k*, *kh*, *g*, *gh*. (The aspiration of NR is quite strong and the aspirated series are interpreted as consonant sequences because of the possibility of infixation in a few words.) My own suspicion is that the aspiration which I was recording for CGR in syllables reflecting voiced stops may have been weaker and perhaps actually a breathiness that I was hearing. It is also possible that the difference between voiced and voiceless that I was recording was also a syllable-level feature that I was not cued into. The one environment in which I recorded less variation was in pretonic syllables where I did not record any aspiration. I suspect that CGR is well advanced in the development of a register system which my NR ear simply was not attuned to and which I did not have time to develop in the short session I had available with linguistically unsophisticated speakers. In examples (1)– (10) I list a number of the words as I recorded them, but elsewhere I list them with an underlined voiceless symbol (*p̱*, *ṯ*, *c̱*, *ḵ*) which should probably be read as representing some kind of register feature.

	CGR	NR	
(1)	at^{h}əi/ad^{h}əi	adəi	younger sibling
(2)	b^{h}aʔ	baʔ	full
(3)	cala:	jala:t	road
(4)	dɪlah	dɪlah	tongue
(5)	hap^{h}au	habəu	ashes
(6)	ʔiduŋ	ʔidũc	nose
(7)	paɣah	barah	to swell
(8)	tanau/danau	danau	pond
(9)	t^{h}lai/d^{h}lai	dlai	forest
(10)	t^{h}ua	dua	two

2.2 SIGNIFICANCE FOR CHAMIC

The apparent developments in CGR of words with original syllable-initial voiced stops would tend to support Thurgood's (1995) thesis that the various developments from proto voiced stops are due to external influence (although in this case the influence is probably not external to Chamic, but is that of Cham itself on CGR). It is not yet clear whether the system of CGR is more like the current status of Eastern or Western Cham, but CGR is not in direct contact with any Mon-Khmer language, so that any external influence would be expected to show the influence of ECham (Eastern Cham) in which many CGR speakers would be bilingual.

Which of the Chamic languages CGR has parallels with in respect to reflexes of words with syllable-initial voiced stops will depend on what the actual manifestation is in current CGR. If voicing has been lost and aspiration predominates, then it is quite parallel to Haroi, which has reanalysed them as aspirated stops (Lee 1977). It should be noted that the pretonic reflexes of voiced stops in CGR are without aspiration and the same is true in Haroi. Haroi does have an occasional allophonic aspiration of voiceless stops in pretonic position, but is unrelated to the source of the stop (Mundhenk and Goschnick 1977:5ff.). It would also be parallel to Utsat, which has voiceless aspirated stops reflecting original voiced stops in the tonic syllable (e.g. Maddieson and Pang 1993).

3. LOSS OF FINAL STOPS FOLLOWING LONG VOWELS: VOWEL LENGTH IN OPEN SYLLABLES

Vowel-length contrast is a feature which is very widespread throughout Southeast Asia. All of the mainland Chamic languages have long and short vowels although the distribution varies somewhat. Rade and Jorai contrast vowel length in all environments except in open syllables and before the laryngeals glottal stop and *h*; Haroi contrasts vowel length in all environments except in open syllables (Mundhenk and Goschnick 1977); Eastern Cham contrasts vowel length in all environments except open syllables and before *h*. NR is the most restricted, contrasting vowel length only before stops including glottal stop. (The national language Vietnamese contrasts vowel length only in closed syllables and that only for central vowels.)

3.1 THE CGR EVIDENCE

For the most part I did not attempt to mark vowel length for CGR except preceding final glottal stop. Consistently where NR has long vowels before oral stops, the final stop was dropped in CGR. Not realising the significance of what was happening, I did not pay attention to the length of the resultant vowel until it became clear that *t^ha* 'one' (NR *sa*) was minimally contrastive with *t^ha:* 'house' (NR *sa:k*). Although I do not have clear evidence that all the long vowels of the words which lost the final stop were retained as long in CGR, there is little reason to doubt that they were. It is possible that CGR contrasts vowel length only in open syllables and before glottal stop, but I could have missed it in other environments where I was not attuned to hearing length contrast.

Words which lose the final stop are listed in examples (11)–(27) along with the NR cognates, and the PC forms are also listed if they have been reconstructed. Clearly long and clearly short vowels in PC are marked as such. Vowels in PC are not marked as long or short when it is uncertain which to reconstruct or when vowel duration is neutral (as in pretonic syllables). It should be noted that most of the words not marked for length are of Austronesian origin but that no contrastive length is reconstructed for them in Proto Austronesian. The CGR forms given below are all marked as long, even though this quality has not been confirmed.

	CGR	NR	PC	
(11)	*pula:*	*bila:t*	**bila:n*	moon
(12)	*cala:*	*jala:t*	**jala:n*	road
(13)	*ṭlɔ:*	*dlɔ:k*	**dlɔ:ŋ*	tall, long
(14)	*ʔdə:*	*ʔdə:p*	-----	to say (ECham *ʔdo:m*)
(15)	*ʔdu:*	*ʔdu:c*	**ʔdo:ŋ*	to float
(16)	*haṭu:*	*hadu:p*	**hadom*	how many (ECham *tɔ̀:m*)
(17)	*huca:*	*huja:t*	**huja:n*	rain
(18)	*ʔika:*	*ʔika:t*	**ʔika:n*	fish
(19)	*krɔ:*	*krɔ:k*	**krɔ:ŋ*	river
(20)	*lupa:*	*luba:k*	**luba:ŋ*	hole
(21)	*pit^ha:*	*pisa:k*	-----	husband (reference in NR) (from **pu* 'lord' + **sa:ŋ* 'house')
(22)	*t^ha:*	*sa:k*	**sa:ŋ*	house
(23)	*t^hia:*	*sia:p*	**sia:m*	good
(24)	*tia:*	*-tia:t*	**tia:n*	abdomen (NR in *matia:t* 'pregnant')
(25)	*tola:*	*tula:k*	**tula:ŋ*	bone
(26)	*ʔu:*	*ʔu:k*	**ʔo:ŋ*	husband (address in NR)
(27)	*wa:*	*ra:c*	------	classifier for people

Words which have final glottal stop following a long vowel, but do not lose the glottal stop, include:

	CGR	NR	PC	
	CGR	NR	PC	
(28)	*ʔayuːʔ*	*ʔayuːʔ*	**ʔayup*	to blow
(29)	*paːʔ*	*paːʔ*	**paːʔ*	four
(30)	*ʔbuːʔ*	*ʔbuːʔ*	**ʔbuk*	hair

Words with short vowel plus oral stop which retain the final stop include:

	CGR	NR	PC	
	CGR	NR	PC	
(31)	*gulap*	*gulap*	**gulăm*	to carry on shoulder
(32)	*ʔbək*	*ʔbək*	**ʔbə̆ŋ*	to eat
(33)	*phut*	*phut*	**phŭn*	trunk/stalk

3.2 SIGNIFICANCE FOR CHAMIC

First, it is possible that CGR is the only Chamic language with a straightforward long/short vowel contrast in open syllables. Fuller (1977) gives a couple of examples with long vowels in open syllables, but the description is sketchy and it is unclear whether there is any glottalisation involved. I do not recognise Roglai cognates for the examples given. There are also examples with long vowel plus glottal stop for which I do recognise cognates. The Chru primer (anon. 1970) has numerous word-final vowels marked with a grave accent, but these vowels are phonemically long and followed by glottal stop. Most of these are cognate with NR and ECham words ending with a long vowel plus glottal stop.

Second, the development in CGR supports my earlier analysis of NR in two respects. Although no acoustic studies have been done, I was aware that usually the NR final stops following long vowels were somewhat lenis. The lenis feature of final [k] following a long vowel was such that Fr Corentin Savary, who spoke Roglai fluently, responded to it as being more similar to his French /r/ than to a [k] although he clearly perceived /k/ following a short vowel as [k]. I even wondered if NR was in the process of beginning to lose these final lenis stops, but I fully expected that the glottal stop was replacing them since I also thought I perceived some accompanying glottal stricture.

I had also analysed word-final vowels in NR as short even though there was no contrast with long vowels. In other languages, analysts had taken the open-syllable vowels to be long. This made a difference in how words were analysed in teaching literacy since in Roglai it made the open-syllable vowel short like the unmarked vowels before stops. Tests done with native speakers tended to support their perception of the open-syllable vowels as being short rather than long even though actually neutralised in the position. The loss of the final consonants in CGR with resulting long open syllables contrasting with short open syllables shows that in CGR the already existing open syllables were short.

4. NASALISATION AND EFFECTS OF NASALISATION

4.1 THE CGR EVIDENCE

Nasalisation of vowels, whether restricted to the vowels or a feature of syllables or words, either occurs or has left its marks in most of the Chamic languages. CGR is no exception. I recorded nasalisation in only a small percentage of the words, but in a much

larger number there is clear evidence of nasalisation. I suspect that, for many of these, nasalisation may still be a significant feature even though I did not perceive or record it. For the words where I did record nasalisation with the vowel, only one was in an environment with no nasal consonant, namely *sə̃p* 'language', which is *săp* in ECham and *sanăp* in NR. It is conceivable that the CGR could be analysed as *səm?* rather than *sə̃p*. In any case it leaves some unexplained details since both ECham and CGR would be expected to have initial aspirated *t* instead of *s*.

What is of interest here is what has happened in CGR to syllables where vowel nasalisation can be reconstructed at least for Lowland Chamic (LC). As Durie (1990:108) has noted, I reconstructed nasalisation for Proto Chamic in cases where I had only the evidence of Roglai. In many cases to do so was probably assuming too much. Unfortunately the reconstructions did not include Haroi or Chru, which would have provided some of the evidence for deciding one way or the other. It is also unfortunate that Jorai, which I did use, has considerable nasalisation, but that this nasalisation was not recorded in materials available at the time. In the data below I have included the ECham cognates, many of which also attest nasalisation, and in a few instances the Chru cognate is also given.

4.1.1 RAISING OF LOW CENTRAL NASALISED VOWEL TO HIGH CENTRAL ARTICULATION

One development from nasalisation in CGR is the raising of the low central nasalised vowel to a high central vowel. The examples (34)–(40) are all between nasal consonants or between a nasal consonant and a final *h* or glottal stop because in other environments there are always further developments which are discussed below. There are also other developments before *h* and glottal stop (see §4.1.3).

	CGR	NR	ECham	
(34)	*lamin*	*lumãn*	*limin*	elephant
(35)	*mɨ̃?*	*mã?*	*mi?*	to get
(36)	*mih*	*mãh*	*mih*	gold
(37)	*mim*	*mãm*	*măm*	to nurse
(38)	*panih*	*panãh*	*panah*	shoot
(39)	*piniŋ*	*pinãŋ*	*paniŋ*	betel nut
(40)	*tanih*	*tanãh*	*tanih*	land

Exceptions are *nam* 'six' and *ŋã?* 'to do' (cf. NR *nãm* and *ŋã?*). ECham *năm* and *ŋa?* do not have the expected reflex of a nasalised vowel. In *panah* 'to shoot' ECham also has *a*, the expected reflex of the low central <u>oral</u> vowel rather than the expected *i*, and Chru has an oral vowel rather than the expected nasal vowel, whereas NR has *ã* and both Western Cham and CGR have *i*, the expected reflex of *ã* for those two languages. The reflexes of 'to suck' are varied throughout Chamic with the following vowels occurring between two *m*'s: *i, ɛ, a, i*.

4.1.2 LOSS OF LOW CENTRAL NASALISED VOWEL ADJACENT TO HIGH CENTRAL VOCOID

If the nasalised low central vowel *ã* is adjacent to a high front vocoid (preceding and/or following), the low central vowel is lost. Examples (41)–(44) are the only ones I have which do not also have a further accompanying change. It is possible that the word for 'right' in CGR is nasalised, but if the word for 'left' had been nasalised, I believe I would have noted that.

	CGR	NR	ECham	
(41)	ʔiu	ʔiãu	ʔiu	left
(42)	hɔnuʔ	hanuãʔ	hnɨʔ	right
(43)	hun	huan	------	mist
(44)	tʰeũːʔ	chiãːʔ	siawʔ	wing (cf. Chru *siãːuʔ*)

It would appear that examples (41)–(44) actually have two changes (raising of *ã* and subsequent loss) which are phonetically conditioned and ordered. Raising of the nasalised vowel from a low central articulation to a high central articulation adjacent to a high front or high back vocoid gives two adjacent vowels which are very similar. The high central vowel, then, was assimilated by the high front or high back. (It should be noted, however, that the national language Vietnamese does have sequences of high central vocoid followed by a high back or high front vocoid.) The loss of nasalisation would have followed the raising of the vowel, but order of the loss of the high central vocoid and nasalisation could have been in either order or simultaneous.

Thus the ordering for CGR in the case of *iu* 'left' could be any one of the following:

(45)   ʔiãu —> ʔɨ̃u —> ʔĩu —> ʔiu

(46)   ʔiãu —> ʔɨ̃u —> ʔĩu —> ʔĩu

(47)   ʔiãu —> ʔɨ̃u —> ʔiu

In the example of 'mist' (43), the NR is not nasalised, but it is assumed that with the shift of earlier *-l* (attested for 'mist' in other Roglai dialects) to *n* in CGR, nasalisation of the vowel also developed and the *ã* subsequently was raised and lost following *u*. In NR the vowel preceding an original *-l* does not nasalise (*ʔuan* 'stuck in the throat' from earlier *ʔual* along with *ʔuãn* 'very' provide the only minimal contrast of oral versus nasal before a nasal consonant in NR.)

The shift of *i* to *e* in *tʰeũːʔ*, the CGR reflex of 'wing' (44), is not explained. It could possibly be related to the preceding voiceless consonant or may have been lowered because of the influence of the earlier low central vowel although it is not lowered in ECham. The ECham form *siawʔ*, however, does not reflect a nasalised vowel although the Chru form *siãːuʔ* is nasalised.

Exceptions recorded where one would have expected loss of *ã* but where it was retained are: *mahãu* 'thirsty' (NR also *mahau*), onomatapoetic *miãu* 'cat' (NR also *miau*), *mãi* 'to come' (NR also *mai*), *nãu* 'to go' (NR also *nau*). Apart from ECham *mahu* 'thirsty', there is no loss in the cognate words in ECham either. The following words in CGR do not reflect the loss following a nasal consonant, but the NR forms have oral vowels, not nasal: *c̲anau* 'pond' (NR *danau*), and *caŋua* 'winnowing basket' (NR *caŋua*).

4.1.3 ACCRETION OF EPENTHETIC VELAR NASAL CONSONANT FOLLOWING NASAL VOWEL

The other significant development is the addition of an epenthetic velar nasal consonant following the nucleus of the syllable. This has been observed for open syllables and with final *h* and glottal stop. Examples (48)–(49) provide the only examples available of the accretion of the nasal with no observed change in the quality of the vowel itself. (The words are aligned in (48)–(58) to make the like features stand out.)

	CGR	NR	ECham	Chru	
(48)	*huŋ*	*hmũ*	*hu*	*hũ*	to have
(49)	*ɲuŋ*	*ɲũ*	*ɲu*	*ɲũ*	he, she, it

Sets (50)–(55) are all examples of the shift of a syllable-final low central vowel to high central vowel plus the accretion of a velar nasal consonant following the vowel.

	CGR	NR	ECham	
(50)	*ʔamɨŋ*	*ʔamã*	*ʔamɨ*	father
(51)	*puwɨŋ*	*buŋã*	*piŋu*	flower
(52)	*hmɨŋ*	*humã*	*hmu*	rice paddy
(53)	*hɨŋ*	*hã*	*hɨ*	you (sing.) (low form in ECham)
(54)	*ləmɨŋ*	*limã*	*limɨ*	five
(55)	*tamɨŋ*	*tamã*	*tamɨ*	enter

Original oral vowels following a nasal consonant do not acquire the velar nasal: *hana* 'to parch' (NR *hana*), *tano* 'male (animal)' (NR *tano*). There were no exceptions in the data available. (The ECham forms for (51)–(52) have *u* instead of *i* as in the other reflexes because these two forms reflect metathesis of pretonic *u* and onset nasal plus application of the same rule as described for CGR in §4.1.2 above.)

Sets (56)–(58) are examples of the accretion of a velar nasal consonant preceding glottal stop and *h*, with the latter being manifested as a voiceless velar nasal, but written here as *ŋh*. (See example (61) in §4.1.4 for one further example preceding glottal stop.)

	CGR	NR	ECham	
(56)	*ʔanɨŋʔ*	*ʔanã:ʔ*	*ʔanɨʔ*	child
(57)	*lumɨŋʔ*	*lumã ʔ*	*limɨʔ*	fat
(58)	*lumɨŋh*	*lumãh*	------	rhinoceros

Exceptions: The reflex of 'rhinoceros' is the only example I have of the acquired voiceless velar nasal, whereas there are three counterexamples (36, 38, 40 above). There are also a couple of exceptions before glottal stop: *laɲi:ʔ* (NR *laɲī:ʔ*) 'sky', *ŋaʔ* 'to do' (NR *ŋãʔ*); for the latter see above in §4.1.1.

Undoubtedly with further study other examples of the accretion of a velar nasal before both *h* and glottal stop would be found. The question arises as to why the CGR data had as many exceptions to the accretion rule as there were. I have already mentioned the high degree of contact between the CGR speaking people and the ECham people. This suggests two possibilities: (1) heavy influence of ECham has caused some of the forms to be borrowed from Cham, or (2) I was getting Cham words instead of CGR. I would like to suggest the

strong possibility of the latter. ECham no longer retains nasalisation as a part of its basic phonemic system. With this information it is reasonable to assume that just as the NR are looked down on by neighbouring language groups because they have strange sounds made through the nose, the CGR speakers also may be looked down on because of *their* strange nasal sounds. Because of the heavy friction through the nostrils, the final voiceless velar nasal sounds especially strange, so that the CGR people could be abandoning it in favour of the Cham cognates or simply avoiding it in situations where they might be looked down on. Most of my data came from two speakers and I do know that the word *lumiɲh* 'rhinoceros' came from the second speaker whereas the three words with *ih* came from the first speaker. Forms like *ʔaniɲʔ* 'child', however, which sounds less strange, were used by both speakers.

4.1.4 LOSS OF LOW CENTRAL NASAL VOWEL AND ACCRETION OF VELAR NASAL CONSONANT

Examples (59)–(61) show the loss of the original low central nasalised vowel *ã* after raising to a high central vowel plus the accretion of the velar nasal. Beginning with (59) some sets of examples have the reconstruction for Proto Lowland Cham (PLC).

	CGR	NR	PLC	
	CGR	**NR**	**PLC**	
(59)	*caniŋ*	*riɲiã*	**ta(ri)ŋã*	ear (cf. ECham *taɲi*)
(60)	*ɟuŋ*	*ɟrãu*	**ɟrãu*	medicine
(61)	*tʰuŋʔ*	*srã:ʔ*	**srã:p*	crossbow (cf. Chru *srã:uʔ*)

Only NR retains a reflex of the Proto Austronesian liquid in the word for 'ear' (59), but most of the Chamic languages show the metathesis of the **i* with the following *ŋ*. It is this metathesis that CGR and ECham reflect with the remaining *i* in the tonic syllable. In the other two examples (60 and 61) it is likely that in CGR *r* is lost before *u*. I suspect that, as in NR, the *r* in clusters in CGR is a high unrounded central vocoid and that the vocoid-loss rule (see §4.1.2) is applied to it even though the origin is different.

The Chru evidence for 'crossbow' (61) is provided to corroborate the source of the CGR *uʔ* from PLC **-p*, for which see §9.1 below.

4.1.5 EVIDENCE OF EFFECTS OF NASAL CONSONANT ACCRETION IN PROCESS

There is some evidence that some of the processes described above have left CGR with unusual reflexes.

One bit of evidence is that the word given for tiger was *ramo* without the expected final *ŋ* of PLC **lumɔ̃:ŋ/rumɔ̃:ŋ*. The form *ramo* would appear to be a back formation with the velar nasal dropped and thus giving evidence of a period when nasal accretion was still an active process in variation with its absence.

Another bit of evidence is the word for 'hand', which has replaced the final alveolar nasal *n* with the velar nasal *ŋ* as in:

	CGR	NR	PLC
	CGR	**NR**	**PLC**
(62)	*tagiŋ*	*taɲãn*	**taŋã:n*

This would appear to reflect a shift from final *n* to *ŋ* during the time when the open-syllable high central nasalised vowel was acquiring a final *ŋ*. Presumably what happened was that during the process, *n* alternated with its absence and then parallel to the other words acquiring *ŋ* it also acquired an *ŋ*.

4.2 SIGNIFICANCE FOR CHAMIC

Although, as Durie has noted (1990:108), nasalisation is somewhat unstable both in Chamic and in Acehnese dialects, there is a thread of consistency which makes it clear that it has been around for a long time.

Of the mainland Chamic languages, only Rade neither has nasalisation nor, to my knowledge, evidence of having had it. I see no reason to believe, however, that Rade did not have nasal vowels even though no trace of it remains. Durie himself (1990:108ff.) indicates that there is some evidence for contrastive nasalisation in Proto Aceh-Chamic (PAC). Furthermore, instability of nasalisation is not a feature unique to Chamic and Acehnese (see Mattisoff 1975:279ff.) and instability of nasalisation is not the only unstable feature of Chamic.

I have not done a systematic study, but it is safe to say that there are a number of features of Chamic which are not stable. For nasalisation Durie cites, for example, the fact that Cham (ECham) has *panah* 'to shoot' rather than the expected **panih* if reflecting a PC nasal vowel, and as I have noted above Chru has *pənah* rather than the expected **pənih*. On the other hand, WCham and CGR both have the expected *pənih*, and NR is *panãh*. (Since Acehnese reflects an earlier oral vowel (Proto Acehnese **panah* instead of **panɔh*), Durie suggests reconstructing Proto Aceh-Chamic **panah* with an oral vowel which was subsequently nasalised in Acehnese after vowel lowering. Durie did not have access to the WCham data.) If we look at various phonological diachronic developments in Chamic, it is almost impossible, if not altogether impossible, to find any single change that does not have several exceptions. I suggest that many of these incomplete changes or exceptions to change are the result of dialect influence. For example, the reasonably consistent change of final nasal consonants to oral stops following oral vowels in all of the languages and/or dialects known as Roglai has some exceptions. NR has *ʔidŭk* reflecting PC **ʔidŭŋ* 'nose', and CGR has *ʔiduŋ* instead of the expected **ʔiduk*. And at least some dialects of Southern Roglai have *prɔŋ* 'big' from PC **prɔŋ*, where one would expect **prɔk* in all forms of Roglai. Because of their geographical spread, with all Roglai dialects except NR being in direct contact with either Chru or ECham, which retain the final nasals, this type of incomplete change is not unexpected.

Having said this, it may be also be safe to say that nasalisation in Chamic is less stable than most other features. What I do not know, and would find helpful to know, is the extent of nasalisation as a feature of neighbouring Mon-Khmer languages. I am aware that it is not a feature of Koho, the only Mon-Khmer language adjacent to NR, and I am unaware of its being a feature of any other non-Chamic language contiguous to the Chamic languages, but that does not rule out the possibility of earlier contact with a language with contrastive nasalisation. It is the speakers of Koho who considered NR to be strange because of the nasal vowels. I also suspect, as I noted above, that the synchronic absence of nasalisation in

ECham has had some influence on the feelings of CGR speakers concerning nasalisation in their language.

4.2.1 PARALLEL DEVELOPMENTS IN CGR AND ECHAM

There is a considerable amount of parallel between the effects of nasalisation in CGR and Cham. I include only ECham here because I have not had a close look at WCham, although I am aware that WCham does share the same evidences of previous nasalisation. Both of the developments described in §4.1.1 and §4.1.2 are shared by CGR and ECham. These are the raising of the low central nasalised vowel to a high central vowel, and the subsequent loss of that vowel adjacent to other high vocoids. Two of the examples from above are repeated here as (63) and (64). In (63) only the raising of the vowel occurs in both ECham and CGR whereas in (64) both languages share both the raising and the subsequent loss.

	CGR	ECham	NR	PLC	
(63)	*lamɨn*	*limɨn*	*lumãn*	**lumãn*	elephant
(64)	*hənuʔ*	*hnŭʔ*	*hanuãʔ*	**hanuãʔ*	right

Although it is possible that Cham and CGR could have independently developed the raising of *ã* to a high central vowel and the subsequent loss of the high vowel contiguous to another high vocoid, it is highly unlikely that the developments were independent, given the proximity of CGR to ECham and the regular contact between the two.

ECham does not share with CGR the accretion of the nasal consonant in open syllables nor before *h* and glottal stop as described in §4.1.3. On the other hand, CGR has one instance of alveolar nasal accretion which is shared with NR, namely *kamɨn* 'we (exclusive)' for which Burnham (1976) has reconstructed PLC **kamɨ* (see §1.2). This is the only instance I know of an accretion of a nasal consonant in NR. The accretion of *n* rather than *ŋ* is likely because of the shared features of the high front vowel and *n* (there are no other instances in the CGR data of nasal accretion following a high front vowel).

Chru has a development that is similar to the raising of the low central nasal vowel, but involves the raising of the mid central nasal vowel [ə̃] to a high central nasal vowel [ɨ̃] (Fuller 1977:83). It is written in Chru as a mid central vowel with a tilde indicating nasalisation.

4.2.2 PARALLEL DEVELOPMENTS IN CGR AND ACEHNESE

CGR and Acehnese share the raising of the low central nasal vowel and the accretion of a velar nasal in open syllables. In CGR and ECham, the vowel is raised to a high central position, whereas in Acehnese it is raised only to a mid central position. In the examples I cite the PAc (Proto Acehnese) as reconstructed by Durie (1990).

Example (65) illustrates the raising of the vowel in CGR and PAc and (66–67) illustrate both the raising of the vowel and the accretion of a velar nasal in the two languages.

	CGR	PAc	
(65)	*tanɨh*	*tanə̃h*	earth
(66)	*ləmɨŋ*	*limə̃ŋ*	five
(67)	*tamɨŋ*	*tamə̃ŋ*	to enter

Durie (1990:111) raises the question as to whether Acehnese shares any features with particular Chamic subgroups. Acehnese shares the raising of the low central nasal vocoid [ã] with both Cham (ECham and WCham) and with CGR, although the Acehnese vowel does not rise as high as the vowel of either CGR or Cham. It shares the accretion of the velar nasal consonant [ŋ] only with CGR. The crucial question is whether the shared features of vowel raising and velar nasal accretion are historically common to Cham, CGR, and Acehnese or whether they are independent developments.

Shift of height and/or perception of height of nasal vowels is well attested. Wright (1975:382) summarises his investigation of perception:

> ...vowel nasalisation is accompanied by an auditory lowering of the vowel, except for the vowels [æ] which rises, [ɑ] which changes very little in quality and [ɔ] which also rises. Although it was originally proposed that such auditory effects would be a function of changes in the frequency of the first formant, this correlation was not observed for the high and back vowels. Rather, perception of vowel height in these areas appears to depend on other, poorly understood parameters of vowel quality, a conclusion also reached by Ladefoged. We may conclude that Ohala's claim is feasible: the auditory facts of vowel perception can provide an explanation for the diachronic tendency to lower that was hypothesised, at least for high and mid vowels...

Acehnese (but not Cham or CGR) lowers the high front and back nasalised vowels (Durie 1990:107ff.), conforming to the perceptual results of Wright's study. Of the three low vowels in Wright's study, two of the nasalised ones are perceived as raised, but the one vowel which is raised in Acehnese and in Cham and CGR is the very one which Wright notes as being perceived with only little change.

Ruhlen (1975:339) notes concerning natural systems of oral and nasal vowels:

> ...although OV's and NV's are often described as having the same absolute vowel height, we may hypothesise that where positional differences do exist there is a universal tendency for high and mid NV's to be lower than their oral partners, while low NV's tend to be higher than their corresponding OV's. Like long vowels, then, NV's tend to be centralised with respect to the OV's though this does not imply, of course, that they are centralised for the same physiological reason.

Although Wright's perceptual studies and Ruhlen's natural system account for some change in height, they do not account for the extent of the lowering of Acehnese high nasal vowels to a low-mid position bypassing the Acehnese high-mid position, unless there were two stages of lowering or the lowering predates the development of the mid-high vowels from high vowels (see Durie 1990:104). Nor do they account for the extent of the raising of the CGR and Cham low central nasal vowel to a high position bypassing the CGR and Cham mid central vowel. Ruhlen's hypothesis does, however, fit well with the raising of the Acehnese low central nasal vowel to a low-mid position. The lowering of the Acehnese vowels is not relevant here since no lowering has been observed in Chamic elsewhere; only the raising of the low central ã is relevant.

Raising of low nasal vowels is not uncommon and could therefore easily have been independent developments in Acehnese on the one hand and Cham and CGR on the other. Nonetheless, it is a feature that is shared and could be historically connected. Since virtually all phonological innovations are natural changes, I don't think we should dismiss the

possibility of an historical connection lightly, but should keep in mind that the degree of raising in Acehnese was much less than in Cham and CGR, which may militate against their being historically connected.

The second feature, that is the accretion of a velar nasal consonant in open syllables, is shared only by Acehnese and CGR although, as noted above in §4.1.3, CGR also has some nasal accretion before *h* and glottal stop. (Acehnese also has at least one instance of an accretion of an alveolar nasal in *jameun* 'formerly' (from Durie 1985). No PC form has been reconstructed, but NR has *jumã* 'formerly'.) The alternation between nasal vowel and *ŋ* is not uncommon either. Ohala (1975:297) in his list of predictions and explanations of nasal sound patterns observes, "The alternation of [ŋ] ≈ ṽ should be more common than the alternation of other nasals with ṽ". By alternation, I assume that Ohala means that [ŋ] can be perceived as, and/or replace, a nasal vowel and vice versa. This being the case, it would not be surprising for a language to have velar nasal accretion following a nasal vowel, with or without loss of nasalisation of the vowel. The question which I am not prepared to answer, but which needs answering, is whether any other language(s) in the area also share(s) the feature of nasal accretion, particularly any Austro-Asiatic language(s). If only CGR and Acehnese share the feature it would seem more likely to have some historical connection.

Whether Acehnese shares developments with specific mainland Chamic languages is a crucial issue, as Durie has correctly noted. If it does not share developments with specific languages, then the separation of Acehnese could predate the break-up of the Chamic group on the mainland of Southeast Asia. If, however, it does share developments with specific languages, then the separation of Acehnese probably does not predate the break-up of the mainland Chamic group. In that case whether one uses the term Achino-Chamic or Chamo-Acehic or Aceh-Chamic seems to me to become a moot question. It would in such a case probably be best simply to broaden the term Chamic to include Acehnese. The exclusion of Acehnese from Chamic appears to have been based more on historical facts concerning the migration(s) (e.g. Cowan 1988) of the Acehnese people to Aceh and the significant differences of Acehnese from mainland Chamic, even though the migration(s) undoubtedly did not precede the fifteenth century AD. Admittedly, Acehnese speakers outnumber all of the speakers of mainland Chamic languages, but this is because of rapid expansion of the Acehnese at the same time some of the other languages may have been becoming depleted. The contribution of CGR toward answering this question is only a small step. As Durie has also noted, a lot of work still needs to be done utilising more accurate data and including older Cham materials, evidence from Utsat of Hainan, and evidence from neighbouring Austro-Asiatic languages. Unfortunately the CGR data itself is scanty and not very accurate, but it does indicate what a wider base can provide.

5. MERGER OF PRETONIC ALVEOLAR AND ALVEOPALATAL STOPS AS ALVEOPALATALS

5.1 THE CGR EVIDENCE

This does not seem to be a consistent change in CGR, but there are a number of words where pretonic **d*- and **t*- have merged with **j*- and **c*-. In the data available there are twice as many exceptions as words which follow the pattern. I suggest that the merger of the alveolar and alveopalatal stops is either a shift in process or an interrupted or reversed shift

resulting from ECham pressure. As already mentioned, it is also possible that I was getting words adjusted toward Cham for my benefit. Examples (68)–(72) all have the expected change although I recorded a variant of 'pond' (72) without the change.

	CGR	NR	PLC	
(68)	cəkuh	tukuh	*tukuh	rat
(69)	caɲiŋ	riŋiã	*ta(ri)ŋã	ear
(70)	cakəi	digəi	*digəi	tooth
(71)	caɣah	darah	*darah	blood
(72)	canau/tanau	danau	*danau	pond

Exceptions noted are:

	CGR	Roglai	PLC	
(73)	tagiŋ	taŋãn	*taŋãn	hand
(74)	takoi	takuai	*takuai	neck
(75)	talai	taləi	*taləi	rope
(76)	tamɨ	tamã	*tamã	enter
(77)	tanɨh	tanãh	*tanãh	earth
(78)	tola:	tula:k	*tula:ŋ	bone
(79)	təha	tuha	*tuha	old
(80)	təke	tukɾi	*tuki	horn (of animal)
(81)	tutʰau	tisəu	*tisəu	breast
(82)	tata	dada	*dada	chest
(83)	tilah	dilah	*dilah	tongue

The following exception shows an unusual k- reflecting *t-:

	CGR	Roglai	PLC	
(84)	kapai	tarapai	*ta(ra)pai	rabbit

5.2 SIGNIFICANCE FOR CHAMIC

It is granted that there has been a steady loss of pretonic syllables and reduction of pretonic consonants in Chamic. The merger of the alveolar and alveopalatal stops in the pretonic syllables in CGR is shared with Haroi where it is very regular. Examples (68)–(72) above are repeated here as (85)–(89) with the Haroi reflexes substituted for the NR reflexes and PC for PLC.

	CGR	Haroi	PC	
(85)	cəkuh	cakɔh	*tukuh	rat
(86)	caɲiŋ	caɲɛa	*ta(ri)ŋa	ear
(87)	caɣah	cariah	*darah	blood
(88)	cakəi	cakhɨ:i	*digəi	tooth
(89)	canau/tanau	caniau	*danau	lake

The set of exceptions (73)–(83) above also all have *c-* in Haroi except that there is no reflex of PC *tanah* 'earth' given for Haroi. The unusual exception in CGR for 'rabbit' (84) is also shared by Haroi (90):

	CGR	Haroi	PC	
(90)	kapai	kapai	*ta(ra)pai	rabbit

The sharing, albeit not thoroughgoing in the CGR data, of the merger of the pretonic alveolar and alveopalatal stops as alveopalatal stops, along with the sharing of the unusual reflexes of PC *ta(ra)pai* 'rabbit', suggest a possible close affinity of CGR and Haroi. Compare Burnham (1976:57ff.) who concluded that Haroi is to be considered as a separate branch of Chamic, that is as neither Highland Chamic nor Lowland Chamic.

The merger of the alveolar and velar pretonic stops in CGR and Haroi differs from the development in Rade where the voiced alveolar and alveopalatal stops fell together with *l-* and *r-* in the pretonic syllable as glottal stop plus *e*, but the voiceless counterparts fell together with *k-* as *k-* (Lee 1966). Examples (85)–(89) are repeated here as (90)–(94) with the Rade reflexes added.

	CGR	Haroi	Rade	PC	
(91)	cəkuh	cakɔh	kkuh	*tukuh	rat
(92)	caɲiŋ	caɲea	kŋa	*ta(ri)ŋa	ear
(93)	caɣah	cariah	ʔerah	*darah	blood
(94)	cakəi	cakhiːi	ʔegei	*digəi	tooth
(95)	canau/tanau	caniau	ʔenau	*danau	pond

Durie (1990:106) has reconstructed for Proto Acehnese at least two instances of pretonic *c-* where PC has *t-*, but in both instances the onset of the tonic syllable is also *c-*, so that it could be assimilation or possibly loss of pretonic syllable with subsequent reduplication.

	PAceh	PC	
(96)	*(cɛ)cɛt	*ticɛ̃ʔ	great-grandchild
(97)	*cucɔ	*ticɔ	grandchild

6. SHIFT OF ONSET *s- TO tʰ-

6.1 THE CGR EVIDENCE

There is a regular shift of *s-* to *tʰ-* in the onset of the tonic syllable and one instance in the pretonic syllable in CGR.

	CGR	NR	PC	
(98)	ʔatʰau	ʔasəu	*ʔasəu	dog
(99)	ʔatʰaʔ	ʔasaʔ	*ʔasap	smoke
(100)	lutʰa	rusa	*rusa	deer
(101)	pitʰa	pisaːk	------	husband
(102)	tatʰiːʔ	tasiːʔ	*tasiʔ	ocean
(103)	tʰa	sa	*sa	one

	CGR	NR	PC	
(104)	tʰaː	saːk	*saːŋ	house
(105)	tʰaʔai	saʔai	*saʔai	elder sibling
(106)	tʰɛūː?	cʰiãː?	*siap	wing (cf. Chru siãːu?)
(107)	tʰia	siaːp	*siaːm	good
(108)	tutʰau	tisəu	*tisəu	breast
(109)	tʰuŋ?	srãː?	*srãːp	crossbow (cf. Chru srãːu?)

Of the three exceptions noted (110–112) only (112) is in the tonic syllable:

	CGR	NR	PC	
(110)	hakʰən	sagə	*sagor	drum
(111)	saɣa	sara	*sara	salt
(112)	săp	sanãp	*săp	speech, language

It was already noted above (§4.1) that ECham also has an initial *s* in *săp* where an aspirated *t* would have been expected from PC **s*. It should be noted that there was only one pretonic instance of **s-* becoming *tʰ-* in CGR whereas there are two where it does not. Further data might give a better pattern, but in both of these exceptions (111–112) CGR also agrees with ECham. For the one CGR word *tʰaʔai* 'elder sibling' where the pretonic **s-* does become *th-*, ECham retains an *s-*. The normal reflex of PC pretonic **s-* in Cham is a simple *h-* as in *hakǎl* 'drum' (grave accent represents low pitch) so that the reflexes with *s-* for both *sara* 'salt' and *saʔai* 'elder sibling' in ECham are unexpected.

6.2 SIGNIFICANCE FOR CHAMIC

To my knowledge only ECham and CGR share the shift of PC **s-* to *tʰ-* and that only in the tonic syllable. I therefore suggest that the shift in CGR is probably because of the contact between the CGR and ECham.

It should also be noted, however, that the two pretonic exceptions noted for CGR and which agree with ECham are also exceptions in Haroi. Like CGR and ECham Haroi has *sara* 'salt' instead of expected **hara*, but *caʔǎi* 'older sibling' instead of expected **ʔaʔǎi*.

7. SHIFT OF -ŋ- AND -r- TO -ɣ- AFTER PRETONIC *a* AND TO -w- AFTER PRETONIC *u*

As with many of the other patterns observed in CGR, this one is not consistent, but happened with sufficient frequency to produce an observable pattern. A larger corpus would, of course, be very helpful.

7.1 THE CGR EVIDENCE

7.1.1 SHIFT OF -n- AND -r- TO -ɣ- AFTER PRETONIC *a*

There are several examples of this pattern. One problem is that for a couple of the words, I recorded *-g-* where I suspect it probably should have been *-ɣ-*. One word I recorded both

ways. In the examples I give them as I recorded them. Northern Roglai does not have the cognate word for 'corn' so I have listed a Southern Roglai form in the NR column.

	CGR	NR	PC	
(113)	ʔagin	ʔanãn	*ʔanan	name (cf. ECham ʔaŋan)
(114)	ʔayiɲ/ʔayin	ʔaɲĩn	*ʔaɲin	wind
(115)	tayəi	təŋəi (SR)	-----	corn (cf. Chru təŋəi)
(116)	tagiɲ	tanãn	*taŋ:n	hand (PLC *taŋã:n)
(117)	paya	bara	*bara	shoulder
(118)	mayiah	mariah	*mariah	red
(119)	saya	sara	*sara	salt
(120)	ṭayah	darah	*darah	blood

The i in CGR ʔagin 'name' is not accounted for since one would have expected i from *ã. ECham, however, also has an i in taɲin 'hand' where i would be expected from PLC *ã. The velar consonant in the tonic onset parallels the Cham ŋ, but I do not know whether SR or Chru also have a velar nasal or not. NR, Haroi, Jorai, and Rade all have n.

No exceptions to the shift of -r- were observed, but a few exceptions were observed for -ŋ-:

	CGR	NR	PLC	
(121)	caɲiɲ	riɲiã	*ta(ri)ŋã	ear
(122)	caɲua	caɲua	*caɲua	winnowing basket
(123)	laɲi:ʔ	laɲĩ:ʔ	*laɲĩ:ʔ	sky

7.1.2 SHIFT OF -n- AND -r- TO -w- AFTER PRETONIC u

In the data available there is only one clear example of each.

	CGR	NR	PC	
(124)	puwiɲ	buɲã	*buɲã	flower
(125)	huwəi	hurəi	*hurəi	day

There is one additional word which I take to be a result of the same change, namely wa 'classifier for people' from PC *ʔuraŋ 'person'. Unfortunately I did not get the word for person, but NR has ʔura:k 'person' and ra:k 'classifier for people'. I would expect CGR to have *ʔuwa:.

A possible counter example noted was ṭayui 'thorn', but this probably derives from an intermediate *daruəi from *durəi with metathesis (as in NR daruəi) and subsequent loss of the ə in CGR. This then follows the expected pattern for the development of *-r- following *a.

7.2 SIGNIFICANCE FOR CHAMIC

The change of -ŋ- and -r- to -w- in CGR is to my knowledge without precedent in Chamic, as is also the change of -ŋ- to -y-. The shift of -r- to the velar fricative -y- is not

without precedent. It also happened in WCham and even the NR has moved in that direction. What I have written as *r* in NR represents a high central non-syllabic vocoid [ɨ] in tonic syllables (*bara* [baɨa] 'shoulder') sometimes with an *r*-like timbre. PC **r-* in pretonic syllables is further weakened in NR and manifested by a lengthening (with two pulses) of the pretonic vowel with no glottal stricture (*radlai* [aadlai] 'Roglai').

I believe that there is probably an historical connection between the NR and the CGR reflexes of PC **r-*. As far as articulation is concerned, the non-syllabic [ɨ] is very close to a voiced velar fricative. The tongue placement is very close and the primary difference seems to be degree of stricture. Whether there is any historical connection between WCham and CGR *y-* as reflexes of PC **r-* is less clear, but, I suggest, should not be ruled out at this stage of our knowledge.

8. SHIFT OF FINAL **-r* TO *-n*

8.1 THE CGR EVIDENCE

A number of words were observed where word final **-r* has become *-n* in CGR. These are:

	CGR	NR	PC	
(126)	*hakən*	*sagə*	**sagOr*	drum
(127)	*kan*	*gə*	**gɔ̆r*	handle
(128)	*pən*	*pə*	**pOr*	to fly
(129)	*piən*	*bhia*	**biər*	short, dwarf
(130)	*ʔutʰan*	*ʔusa*	**ʔusăr*	seed, flesh
(131)	*wan*	*wa*	**war*	pen, cage

The only possible exception noted was CGR *pu* 'cooked rice', but the cognate forms for other Lowland Chamic languages and Haroi also lack any overt reflex of a final **-r*. NR regularly loses **-r*, so all of the NR reflexes in sets (126)–(131) as well as *bu* 'cooked rice' end with a vowel.

8.2 SIGNIFICANCE FOR CHAMIC

The shift of final **-r* to *-n* in CGR is to my knowledge unique in Chamic. It is highly unlikely, however, that the shift from **-r* to *-n* represents a single change. It is probably safe to assume that **-r* changed first to an intermediate **-l* which in turn changed to *-n*. If this assumption is true, CGR shares the first change with Haroi in which PC **-r* is reflected by *-l*.

	CGR	Haroi	PC	
(132)	*hakən*	*ʔakhuːal*	**sagOr*	drum
(133)	*kan*	*khuːɔl*	**gɔ̆r*	handle

	CGR	Haroi	PC	
(134)	pən	pɔːl	*pOr	to fly
(135)	ʔutʰan	ʔasal	*ʔusăr	seed, flesh
(136)	wan	waːl	*war	pen, cage

Whether this shared feature of Haroi and CGR are part of a shared history remains to be determined. Both the alternation of *r* with *l* and of *l* with *n* are historically common (see Ohala 1975:296 for *l* and *n*). The alternation of *r* and *l* is common in Chamic in onset position and NR shares with CGR the reflection of final PC *-l* as -*n* although there is only one instance in the CGR data available:

	CGR	NR	PC	
(137)	kapan	kapan	*kapaːl	thick

9. SHIFT OF FINAL PC *-p TO -uʔ

9.1 THE CGR EVIDENCE

	CGR	NR	PC	
(138)	ʔayuː ʔ	ʔayuː ʔ	*ʔayup	to blow
(139)	huṭiuɋ/həṭi ʔ	hadiu ʔ	*hadĭp	to live
(140)	ʔjau ʔ	ʔjə ʔ	*ʔjŏp	correct
(141)	patiaːu ʔ (to hunt)	tiaː ʔ	*tiəːp	to pursue
(142)	theū ʔ	chiāː ʔ	*siap	wing
(143)	yaːu ʔ	yaː ʔ	*yaːp	to count

In all of the examples (138)–(143) both CGR and NR lose the final *-p* but differ in what happens. CGR retains both the labial feature and the stop feature with the -*uʔ* with the exception of the variant in (139) where the labial feature is lost following *i*. In (138) there is only one *u* before the glottal stop, the *u* from the *-p* having merged with the nucleus. NR normally retains only the stop feature but has one exception (139) where the labial feature is also retained following the *i* in *hadiuʔ.*

The only exceptions noted for CGR in the data available were:

	CGR	NR	PC	
(144)	ʔasaː ʔ	ʔasaː ʔ	*ʔasăp	smoke
(145)	pa ʔ	ba ʔ	*băp	full

9.2 SIGNIFICANCE FOR CHAMIC

The development of *-p* in CGR is shared by both ECham, Chru, and Haroi. I have evidence for only a few of the Chru reflexes. For all the examples above (138–143) for which I have evidence, ECham and Haroi have replaced the *-p* with a labial semivowel and glottal stop except for Haroi *ʔathip* 'alive' which retains the -*p* of *hadĭp.* For the two

exceptions given above for CGR (144–145), the Haroi reflexes have the expected change, but ECham agrees with CGR for both as shown in (146–147):

	CGR	ECham	Haroi	PC	
(146)	ʔatʰa:ʔ	ʔatʰẵʔ	ʔasauʔ	*ʔasăp	smoke
(147)	paʔ	pẵʔ	phɨauʔ	*băp	full

It should be noted concerning these two exceptions they are the only two for which a short ă is reconstructed for PC. Whether this is significant remains to be seen.

10. *puwɨŋ* 'FLOWER' AND *tʰuŋʔ* 'CROSSBOW' AS REFLEXES OF PC

I began with CGR and NR reflexes of PC *buŋa* 'flower' and PC *srã:p* 'crossbow', noting that although CGR and NR are closely related, these two sets of words would make them appear to be quite different. Apart from the change of *-p* to glottal stop in the word for crossbow, the NR reflexes of these two words are structurally unchanged. The CGR reflexes, however, are radically restructured.

CGR *puwɨŋ* 'flower' manifests four and possibly five changes:

1) Change of initial voiced consonant *b* to a voiceless or partially voiceless sound (symbolised by *p* with other possible changes within the word as part of a Mon-Khmer type register system (see §2)).

2) Change of *-ŋ-* to *w* following *u* (see §7.1.2).

3) Change of *-ã* to *ɨ* (see §4.1.1).

4) Accretion of *ŋ* following *ɨ* (see §4.1.3).

5) Possible loss of nasalisation of *ɨ*. Although I did not record *i* as being nasalised, it may have been.

The third and fourth changes dealing with nasalisation as stated above are ordered, and if the fifth change applies it must follow the third and fourth changes. It is possible, however, that nasal accretion preceded the raising of the vowel. Otherwise the order of the changes does not seem relevant. Of the original phonemes of *buŋã*, only the *u* of the pretonic syllable appears unchanged, but even it may have acquired some change of quality as part of a developing register system.

CGR *tʰu:ŋʔ* 'crossbow' manifests six and possibly seven changes:

1) Change of initial *s-* to *tʰ-* (see §6.1).

2) Change of final *-p* to *-uʔ* (see §9.1).

3) Raising of *-ã-* to intermediate *ɨ* (see §4.1.1).

4) Loss of *ɨ* preceding *-u* resulting in *ū:* (see §4.1.2).

5) Accretion of nasal consonant following *ū:* resulting in final nasal consonant followed by glottal stop (see §4.1.3).

6)   Possible loss of nasalisation of *ūːr*.

7)   Loss of *r* as second member of cluster before *u* (see discussion under §4.1.4).

Again, as they are stated, changes (3)–(6) dealing with nasalisation are ordered, and if (6) applies it must follow (2)–(5). Changes (3)–(5) could, however, be reordered if stated differently.

11. SUMMARY OF CGR INNOVATIONS SHARED WITH OTHER CHAMIC LANGUAGES

Throughout the paper, I have indicated features shared with various of the Chamic languages and with Acehnese. Here I list the features shared with each of the languages included which will enable us to see something of the relationship of CGR to the others.

Northern Roglai (See also Roglai)

1)   Share [dl] as opposed to [gl] shared by other Lowland Chamic languages (§1.2).

2)   Accretion of final alveolar nasal consonant in *kamīn* 'we (exclusive)' (§1.2; §4.2.1).

3)   Lexical items: reflexes of *buːr* mean 'cooked rice' (also shared with Acehnese (§1.2)).

4)   NR may be in the process of losing final stops after long vowels with possible long/short contrast in open syllables as in CGR (§3.2).

5)   NR *r* becomes high central unrounded non-syllabic vocoid (except in pretonic syllable). This is probably related to the CGR shift to voiced velar fricative (§7.2). As a second member of a cluster, both tend to lose the *r* preceding a high vocoid, suggesting that *r* first became a high non-syllabic vocoid (§4.1.4).

Roglai (Southern and Northern)

1)   Denasalisation of final nasal consonants and becoming voiceless stops (§1.2).

ECham (some shared also with WCham)

1)   Development of register system connected with original initial stops (§2.2).

2)   Raising of low central nasal vowel to high central position (§4.1.1).

3)   Loss of high central nasal vowel adjacent to high front or back vocoid (§4.2).

4)   Shift of initial *s*- to *tʰ* (§6.2). Share most exceptions as well (§6.2).

5)   Final *-p* becomes final high central non-syllabic vocoid combined with glottal stop (§9.2). (This feature is also shared with Haroi.)

WCham

1) Change of *r- to voiced velar fricative (§7.2).

2) See also ECham although not all features shared.

Haroi

1) Development of register system connected with original initial stops. CGR in development stage and Haroi has already restructured vowel system with accompanying loss of register system (§2.2).

2) Merger of initial alveolar and palatal stops although sporadic in CGR (§5.2).

3) Shared initial *k*- in word for 'rabbit' (§5.2).

4) Final *-r becomes -*l* although in CGR the resulting -*l* along with already existing -*l* became a final -*n* (§8.2).

5) Final *-p becomes final high central non-syllabic vocoid combined with glottal stop (§9.2). (This feature is also shared with ECham.)

Acehnese

1) Raising of low central nasal vowel: CGR to high position, Acehnese to mid-low position (§4.2).

2) Accretion of velar nasal consonant following nasal vowel in open syllables (§4.2).

3) Possibly shift of *t to c (see also for Haroi above), but doubtful (§5.2).

Looking only at the above list, the features shared by CGR with NR, ECham, and Haroi are the highest and roughly equal in number. The number of features shared with some of the other languages may turn out to be as many as when we have the information organised. The nature of some of the features shared with ECham and Haroi, however, seem to indicate that there may be a closer affinity of Haroi with other Lowland Chamic languages than Burnham (1976) recognised. Southern Roglai and also Chru should have a lot in common with CGR that is not available here. I would expect that features shared specifically with Rade or Jorai, however, are somewhat unlikely.

Concerning the relationship of CGR with NR and ECham, I expect that the features shared by CGR and NR are more likely held over from earlier common history and that some of the features shared with Cham are from more recent close contact.

CGR shares two nasal related innovations with Acehnese but the significance of these remains to be seen.

REFERENCES

Aday Bach Akhar Cham Birau: Tapuk 1, 3 [I learn to read New Cham: Books 1, 3] 1970, Saigon: Ministry of Education.

Adoi Posram Ponuai Chru: Top 1-3 [I learn the Chru language: Books 1-3] 1972, Saigon: Ministry of Education.

Burnham, Eugene Courtney, 1976, The place of Haroi in the Chamic languages. MA thesis, University of Texas at Arlington.

Durie, Mark, 1985, *A grammar of Acehnese on the basis of a dialect of North Aceh.* Dordrecht: Foris Publications.

    1990, Proto-Chamic and Acehnese mid vowels: towards Proto-Aceh Chamic. *BSOAS* 53/1:100–114.

Ferguson, Charles A., Larry M. Hyman and John J. Ohala, eds, 1975, *Nasálfest: papers from a symposium on nasals and nasalization.* Stanford, California; Language Universals Project, Department of Linguistics, Stanford University.

Fuller, Eugene, 1977, Chru phonemes. In Thomas, Lee and Liêm, eds 1977:77–86.

Lee, Ernest Wilson, 1966, Proto-Chamic phonologic word and vocabulary. PhD dissertation, Indiana University.

    1977, Devoicing, aspiration, and vowel split in Haroi: evidence for register (contrastive tongue-root position). In Thomas, Lee and Liêm, eds 1977:87–104.

Maddieson, Ian and Keng-Fong Pang, 1993, Tone in Utsat. In Jerold A. Edmondson and Kenneth J. Gregerson, eds *Tonality in Austronesian languages.* Special Publications of Oceanic Linguistics 24:75–89.

Matisoff, James A., 1975, Rhinoglottophilia: the mysterious connection between nasality and glottality. In Ferguson, Hyman and Ohala, eds 1975:265–288.

Mundhenk, Alice Tegenfeldt and Hella Goschnick, 1977, Haroi phonemes. In Thomas, Lee, and Liêm, eds 1977:1–15.

Ohala, John, 1975, Phonetic explanations for nasal sound patterns. In Ferguson, Hyman, and Ohala, eds 1975:289–316.

Ruhlen, Merritt, 1975, Patterning of nasal vowels. In Ferguson, Hyman and Ohala, eds 1975:333–351.

Thomas, David, Ernest W. Lee and Nguyễn Đăng Liêm, eds 1977, *Papers in South East Asian Linguistics No. 4: Chamic Studies.* PL, A–48.

Thurgood, Graham, 1995, Language contact and the directionality of internal 'drift': the development of tones and registers in Chamic. MS.

Wright, James, 1975, Effects of vowel nasalization on the perception of vowel height. In Ferguson, Hyman, and Ohala, eds 1977:373–387.

ON THE ETHNONYM 'UTSAT'

KENG-FONG PANG

1. ETHNONYMS FOR THE UTSAT PEOPLE[1]

Utsat is the ethnonym preferred by the Chamic-speaking Muslims of Hainan Island and can be said to be an autonym.[2] The Utsat now number about 6,000 and are concentrated in two villages near Sanya on the southern tip of Hainan Island. They use *Utsat* to refer to themselves when speaking in their indigenous language, and they refer to their language as *Tsat*. Thus non-Tsat speakers, including those scholars who study the Utsat people without learning to use Tsat, will usually not be aware of this term (see Pang 1992). In English-language literature since 1992 Western linguists and scholars have used the name Utsat (see Pang 1992, Maddieson & Pang 1993, and Thurgood 1993).

The Utsat have been previously known to the world as *Hui*. Hui is the Utsat's ethnic identity or nationality name officially decided upon during the 1950s when the People's Republic of China's Minority Nationalities Commission began work to identify the minority nationalities. Hui, however, is also a Han Chinese term commonly understood to refer to the Chinese-speaking Muslims (see Pillsbury 1989, 1973, and Gladney 1989, 1991), and is often erroneously extended in Han conversation to refer to all Muslims in general.[3] In Chinese linguistic literature, the Utsat have been referred to as Hainan Hui and their language as *Hui-Hui hua* 'Hui-Hui language' (Ouyang and Zheng 1983, Zheng 1986, and Ni Dabai 1988). The local Han people may also refer to the Tsat language as *Hui hua* or possibly *Hui-Hui hua*. But the Utsat people do not normally use *Hui-Hui hua*.

---

[2]   See Benedict (1987) for a discussion of autonyms and exonyms. Benedict (1941) was the first to suggest a Chamic origin for the Utsat as described by Stubel (1937). See Pang (in progress) for a reassessment of the linguistic, archaeological, and cultural evidence regarding the Chamic origin of the Utsat. For Chinese writings on the origins of the Hainan Hui see Dong (1985), Li and Tian (1986), Jiang and Mei (1986), and Chen and Jiang (1988).

[3]   I have argued elsewhere (Pang 1987, 1992:29–38) that the Utsat could have qualified as the eleventh Muslim minority nationality in the People's Republic of China because they are historically, linguistically, and culturally clearly distinct from mainland Hui by every criterion listed by the Minority Nationalities Commission. It is noteworthy that in a graduation essay written by two Utsat religious scholars they chose to call themselves Hainan Muslims rather than Hui.

In speaking Hainanese the Utsat tend to refer to their language as *Huan-uei* 'Huan language', and they themselves are known in Hainanese as the Huan people.

Since not all people on the mainland who identify themselves as Hui are practising Muslims, the more sophisticated Hui and Hui scholars (see Gladney 1991) on the mainland make a distinction between 'Hui' as an ethnic identity and 'Hui' for Muslim or Islam. For the same reason the more sophisticated and better travelled Utsat also prefer to call themselves *Hainan Mu-si-lin* ('Hainan Muslim' in Mandarin) instead of Hui. The Utsat who are not practising Muslims are referred to as having 'become Han'. The term *I-si-lan-jiao* 'Islam religion' is more properly used to refer to the religion of Islam.

When the Utsat emigrated to Malaysia they became known as Orang Kwangtung by the Muslim Malays (see Pang 1994). With the passing of the pioneering generation in Malaysia both the Tsat language and the ethnonym Utsat have dropped out of use.[4]

The social reality in Hainan is complex. My anthropological research among the Utsat for almost two years (1987–1989 and shorter visits in 1991, 1993, and 1994) using a combination of Tsat, Hainanese, Mandarin, and Malay languages exposed me to the fact that the Utsat have multiple ethnic identities. Variously known as Hui, Utsat, and Huan-nang, each ethnic identity is mediated through a specific language use (respectively Mandarin, Tsat, and Hainanese). Each identity has a specific contextualised local history which emphasises different aspects of being Utsat. I have elaborated elsewhere on Utsat's repertoire of simultaneous ethnic identities (see Pang 1995).

In this article, I will focus specifically on the various meanings of the ethnonym 'Utsat' as opposed to other ethnonyms, examining how the ethnonym is used linguistically, and offering culturally grounded analyses of everyday Utsat social interactions and their discourses about being Utsat. I cite several instances where the use of the Tsat language has helped me to understand the term *Utsat* in its multiple meanings and contextual usage.

2. UTSAT AS DIFFERENTIATED FROM OTHER PEOPLES

How is the term 'Utsat' used in everyday discourse? When a stranger walks into an Utsat village, the Utsat people might ask each other the following question:

Nau si Utsat ahsi Ulo?[5]

(He/She is Utsat or Ulo?)

Is he/she an Utsat or Ulo?

On my first day of language-learning through social immersion, this very question alerted me to the existence of the term 'Utsat' which up to that time had not been mentioned in either the Chinese or foreign academic literature. Utsat was clearly being used as an ethnonym/autonym that contrasted with Ulo.

---

4   See Pang (1989) for an analysis of the Utsat's Southeast Asian connections.

5   I am not using the IPA phonetic symbols or making a precise phonetic or phonemic transcription. As an anthropologist who is thus far the only scholar to have done long-term fieldwork using this Austronesian language, I hope this contribution will clarify why the Utsat of Hainan Island are not simply 'Hui'—the term by which they and many Muslims of mainland China who speak Chinese languages as their indigenous languages are known.

Who is an Ulo? In its most inclusive and general meaning, an Ulo is simply a non-believer in Islam, a kafir. This meaning will become clearer when we later discuss the use of Utsat to mean simply Muslim—the explanation most commonly asserted by the Utsat themselves. However, in the everyday local context, Ulo more specifically refers to the Han Chinese who are non-Muslims. Interestingly, the term *lo* in Tsat means 'meat' in general, as in *lo-phui* 'meat-pig' (pork) and *lo-mo* 'meat-cow' (beef). It is not inconceivable that the Utsat first used the term Ulo to refer to non-Muslims in their midst who routinely ate pork, which is abhorred by Muslims. This reasoning finds further support in another Chamic language when we observe that the Western Cham speakers use the term *lo* (in a lower tone than that used in tonal Utsat) to describe the Chinese in Cambodia.[6] Whether this is a Proto Chamic term for both 'meat' and 'Chinese' remains to be seen.

How do these terms, Utsat and Ulo, fit into basic Utsat taxonomic classification of peoples? In their least elaborated taxonomic classification, if asked "How many types/kinds of people are there?", many Utsat would respond first with "Utsat, Ulo, Ulait, and Ulait-miao" before elaborating on each category. It is noteworthy that while the terms Utsat and Ulo might have theoretically covered all the world's people as 'Muslim' and 'Non-Muslim', the Utsat clearly differentiate themselves and the Ulo from the local Ulait and Ulait-miao. The latter two terms are ethnonyms for the Li and Miao people who are considered by Utsat to be "people who stay in the forests". Although many Li people now live in the cities and some Li girls have recently been employed by Utsat as live-in nanny-housekeepers, many Utsat elders recall seeing bare-breasted Li women and Li men in loincloths as recently as 60 years ago. It is also conceptually interesting that the Miao are classified as a subgroup of the Li, even though the Utsat can describe the cultural and linguistic differences between them. The Miao, actually Yao-speakers not linguistically or culturally related to the Miao (Hmong) from mainland China (Jacques Lemoine pers.comm.), are also more feared and less encountered by the coastal living Utsat. I would suggest that, to the Utsat, the Li and the Miao are not only distinct from the Han Chinese but were also viewed in earlier times as being somewhat savage and subhuman because they were forest or hill dwellers. Thus the basic taxonomy includes Li and Miao as separate categories of people.

3. UTSAT AS BEING MUSLIM

Who is an Utsat then? When asked to articulate how Utsat are different from the Ulo, the most commonly listed characteristics were: "We Utsat believe in Allah, the Ulo pray to many gods" or "We Utsat do not eat pork, the Ulo do". Many Utsat also feel a sense of moral superiority over the Han Chinese, a sense which derives from knowing that good Utsat will enter heaven upon death if they practise Islamic teachings well. This sense of moral superiority finds expression among Utsat women sellers of vegetables and fruits in the city markets or by the roadside when they say to me in Tsat (or in Hainanese and Mandarin to Han Chinese):

> *Mi Utsat pu phian dzat.*   'We Utsat do not cheat people.'

The unwavering belief that an Utsat vegetable seller would not short-change her customer in the market minimises price-negotiating because the Utsat will often tell the customer "You do not have to worry, we do not cheat people. Our Allah will know if we cheat".

This Islamic core of Utsat identity is clearly so basic to the Utsat that any Utsat will articulate that 'Utsat' means 'Muslim'. Indeed, to them all the world's Muslims are Utsat people, differentiated only by nationality or region. In Tsat, the practising Hui Muslims on the mainland are described as 'Utsat Talok' (mainland Muslims). The Uighur or other Turkic-speaking Muslims from Xinjiang province would be called 'Utsat Sinkiang'. Similarly, American Muslims would be known as 'Utsat Meikok' (Muslim Americans) while Saudi Muslims are 'Utsat Saute'.

Indeed, we can see how deeply entrenched this Islamic core is by noting the fact that to study Islam or the Koran is usually referred to as *phai kha:d tsat* (to study Arabic/Islamic words). To go to Islamic school is *nauk hok kha:d tsat* as opposed to going to a Han Chinese school which is *nauk hok kha:d lo*. Similarly, each Utsat person typically has three types of names, a Muslim name, a school-going Chinese name, and a nickname. The Muslim name is referred to as *nan kha:d tsat* and is the first name given to a child, usually nine days after birth. Thus the term *kha:d tsat* can be glossed as 'Islamic words' or 'Arabic or Koranic language'.

The Islamic core of Utsat identity is also underscored by the fact that 'becoming Utsat' (*ngau Utsat*) is something to be achieved performatively by first becoming Muslim. Learning the Tsat language comes later. It is unlike the Hui identity, where one is Hui because one has Hui blood (see Pillsbury 1976). Being Utsat and remaining Utsat is a performative act. An Utsat who no longer practises Islam will be said to have become Han (*ngau Ulo*), even if this person still speaks the Tsat language. The converse is also true. Take my position as an anthropologist in the community, for example. I have often been asked to become Utsat for several reasons. Most Utsat say that I should become Utsat because I have lived with them for a long time, because I understand Utsat culture and Muslim customs, and, last, because I already speak Tsat (which means I have the communicative skills to live meaningfully in their society). However, I cannot be called an Utsat unless I decide to convert ('enter the religion') and become a Muslim. Thus, for example, I have been exhorted in Tsat:

> *Ha ma kiau ngau utsat.*   'You enter religion and become Utsat.'

The fact that the Utsat regard the Islamic core of their identity as being central to their articulations of selfhood as a people is not surprising when we note in their myths of origin that only the Utsat who decided to remain living together in the same place with fellow Utsat remained Utsat. Those who decided to stay in Han-dominated areas presumably became Han. Indeed, there are several coastal cities in Hainan where a section of town is known as 'fan-cun' (barbarian or foreign villages) in the historical literature (see Pang 1992). It is noteworthy that the Utsat have no folk or oral history indicating a conversion to Islam, which suggests that the Utsat either came to Hainan Island as Muslims, or were converted too long ago to retain this fact even in their oral tradition. The latter scenario is not inconceivable when we consider that historically there was an ancient Muslim settlement in Hainan which was described by a Chinese traveller in the eighth century as 'Persian' (see Gerini 1974:471, fn.2). The existence of this ancient Muslim settlement might explain who were the Muslim people buried in an extensive area near Lingshui marked with coral Muslim tombstones (see Zhong 1984, Li and Wang 1987, and Chen and Salmon n.d.). It is important to note that

tombstones were not considered by the Utsat to be part of their historical or oral traditions, nor is this area regarded as their ancestral burial ground.

4. UTSAT IN LINGUISTIC PERSPECTIVE

Nevertheless, from both anthropological and linguistic perspectives, it is clear that *Utsat* is meaningful as an ethnonym in ways other than those articulated by the Utsat in everyday discourse. For example, when we see the linguistic correspondence between the terms Utsat, Ulo, and Ulait whereby Utsat people speak Tsat, Ulo people speak Lo, and the Ulait people speak Lait, it is reasonable to infer that the term Utsat means 'people who speak Tsat'. The prefix 'u' in each of these ethnonyms is clearly a root Austronesian term which means 'people'. This is, however, a meaning of Utsat which is not articulated by the Utsat themselves. This suggests that the morpheme 'u' is no longer a productive morpheme in modern Tsat language (Maddieson, pers.comm. 1992).

From a comparative Chamic perspective, an etymology of the term 'Utsat' has been offered by Mark Durie (pers.comm. 1993) who suggested that the term *Tsat* corresponds regularly with *Cham*, with the loss of labial articulation in the final nasal of *Cham*.

5. CONCLUSION

In analysing the various meanings of *Utsat* among the Utsat people, there is clearly a conflation of both religious and cultural elements, with the Islamic core of their identity superseding all others. It is not enough to be able to speak the Tsat language to be Utsat, one must be a practising Muslim.

REFERENCES

Benedict, Paul R., 1941, A Cham colony on the island of Hainan. *Harvard Journal of Asiatic Studies* 6:129–134.

1987, Autonyms: ought or ought not. *Linguistics of the Tibeto-Burman Area* 10.1:186–188.

Chen Dao Liang and Jiang Qing Wu, 1988, Hai Nan Mu Si Lin [Hainan Muslims]. Graduation Paper, Beijing Centre for Islamic Studies.

Chen Dasheng and Claudine Salmon, n.d., Rapport préliminaire sur la découverte de tombes musulmanes dans l'île de Hainan. Unpublished manuscript.

Dong Zhen Qi, 1985, Hai Nan Dao Hui Zu Zu Yuan Wen Ti Chu Tan (On the origin of Hui nationality on Hainan Island). MA thesis, Department of Anthropology, Zhongshan University.

Gerini, G.E., 1974 (1909), *Researches on Ptolemy's geography of eastern Asia (Further India and Indo-Malay Archipelago)*. London: Asiatic Societies Monographs, No. 1.

Gladney, Dru C., 1989, Identifying the Hui in China: self-definition, ethnoreligious identity, and the state. In Gladney, comp. 1989. *The legacy of Islam in China: an international symposium in memory of Joseph F. Fletcher*. Conference volume, Harvard University, April 14–16, 1989.

Gladney, Dru C., comp., 1989, *The legacy of Islam in China: an international symposium in memory of Joseph F. Fletcher*. Conference volume, Harvard University, April 14–16, 1989.

1991, *Muslim Chinese: ethnic nationalism in the People's Republic of China*. Cambridge: Harvard East Asian Monographs No. 149.

Jiang Yong Qing and Mei Wei Lan, 1986, Hainan Dao Yanglan Huizu De Lai Yuan Ji Qi Te Dian [The origins and characteristics of Yanglan Hui nationality in Hainan Island]. *Ningxia Social Sciences* 4:21–26.

Li Qi Li and Wang Ke Rong, 1987, A view of ancient Muslim activities on Hainan Island from the relics in Linshui Sanya Muslim tombs. *Hainan Museum Journal* 1:53–59.

Li Ce Jin and Tian Guang, 1986, Hainan Dao Yanglan Huizu De Li Shi Ji Xian Zhuang Diao Zha [The history and current investigation of the Hui Nationality in Yanglan, Hainan Island]. *Minzu Yan Jiu Dong Tai* [Current Nationality Research] 4:26–31.

Maddieson, Ian, and Keng-Fong Pang, 1993, Tone in Utsat. In Jerold A. Edmondson and Kenneth J. Gregerson, eds *Tonality in Austronesian languages*, 75–89. Special Publications of Oceanic Linguistics No. 24.

Ni Dabai, 1988, Hainan Dao Sanya Huizu Yuyan de Xishu [The genealogical affiliation of the language of the Hui people in Sanya, Hainan]. *Minzu Yuwen* 2:18–25.

Ouyang, Yueya and Zheng Yi Qing, 1983, Hai Nan Dao Ya Xian Hui Huizu de Hui Hui Hua [The Hui Hui Language of the Hui nationality in Yaxian Hainan]. *Minzu Yuwen* 1:18–30.

Pang, Keng-Fong, 1987, The Austronesian-speaking Muslims of Hainan—Utsat or Hui? A problem in the People's Republic of China's identification of minority nationalities. Paper presented at the Meeting of Southwestern Anthropological Association, Bakersfield, March 26–28.

1989, The Muslims of Hainan and their Southeast Asian connections: a preliminary analysis. In Dru Gladney, comp. 1989.

1992, The dynamics of gender, ethnicity, and the state among the Austronesian-speaking Muslims (Hui/Utsat) of Hainan Island, People's Republic of China. PhD dissertation in Anthropology, University of California, Los Angeles. Ann Arbor: University Microfilms International.

1994, How Orang Kwangtung from China became "Melayu": the "Malayization" process from four communal perspectives. Paper presented at the 46th Annual Meeting of the Asian Studies Association, Boston, March 24–27.

1995, Being Hui, Huan-nang, and Utsat simultaneously: contextualizing history and identities of the Austronesian-speaking Hainan Muslims. In Melissa Brown, ed. *Negotiating ethnicities in China and Taiwan*, 183–207. Berkeley: Institute of East Asian Studies.

in progress, Paul Benedict's "A Cham colony in Hainan Island" revisited: assessing the new evidence for the Chamic origins of the Utsat (Hui).

Pillsbury, Barbara L.K., 1973, Cohesion and cleavage in a Chinese Muslim minority. PhD dissertation, Department of Anthropology, Columbia University.

1976, Blood ethnicity: maintenance of Muslim identity in Taiwan. Paper presented at Conference on Taiwan, Portsmouth, England, August 19–24, 1976.

1989, Transformation of Hui identity in Taiwan and China. In Dru Gladney, comp. 1989.

Stubel, H., 1937, *Die Li-Stämme der Insel Hainan: ein Beitrag zur Volkskunde Südchinas, unter Mitwirkung von P. Merriggi*. Bern, Klinkhardt and Biermann.

Thurgood, Graham, 1993, Phan Rang Cham and Utsat: tonogenetic themes and variants. In Jerold A. Edmondson and Kenneth J. Gregerson, eds *Tonality in Austronesian languages,* 91–106. Special Publications of Oceanic Linguistics No. 24, University of Hawaii Press.

Zheng Yi Qing, 1986, Zai Tan Hui Hui Hua de Di Wei Wenti [A further discussion of the position of Huihui speech and its genetic relationship]. *Minzu Yuwen* 6:37–42.

Zhong Yuan Xiu, 1984, Hai Nan Dao Yi Si Lan Gu Mu Qun Kao Cha Jian Ji [A brief investigation recording ancient Muslim tombs in Hainan Island]. *Minzu Yan Jiu Dong Tai*. Current Nationality Studies No. 4.

AUSTRONESIAN AND MON-KHMER COMPONENTS IN THE PROTO CHAMIC VOWEL SYSTEM

GRAHAM THURGOOD

1. INTRODUCTION[1]

The Austronesian speakers who arrived on the coast of the Southeast Asian mainland spoke a basically disyllabic language with a relatively modest vowel inventory. The morphemes were typically disyllabic, more specifically, CVCV(C), and there were four basic vowels: *-a, *-i, *-u, *-e (= [-ə]) and three final diphthongs: *-ay, *-ui, and *-aw; the four vowels occurred in both syllables of the disyllabic forms while the diphthongs were restricted to the final syllable.

Under the influence of what was apparently more than casual contact with Mon-Khmer (MK) languages, this pre-Chamic Austronesian (An) language adopted the main-syllable stress of the neighbouring MK languages, a change that had consequences both for the morpheme structure and for the vowel inventories of Proto Chamic (PC). By the time of PC, the formerly disyllabic Austronesian roots had become iambic (in the sense of Donegan 1993); that is, the formerly disyllabic morphemes came to have an unstressed initial syllable followed by a stressed main syllable. This iambic PC stress pattern is certainly reflected in the contrasts between the vowel inventories of the pre-syllable and the main syllable. Unlike in the Austronesian disyllables where there was a balanced four-way vowel contrast in both syllables,[2] in PC the vowel inventories are anything but symmetrical: in the unstressed PC pretonic syllable, the four-way Austronesian vowel distinction has been reduced to a three-way distinction while in the stressed main syllable the same four-way distinction, has been expanded to 18 or so distinct vowels, not counting length contrasts.[3]

Some of these new main-syllable vowels developed out of splits of inherited Austronesian vowels, but the bulk of the forms with new vowels are found in pre-Chamic borrowings from MK. Thus, the main vowels of PC include two readily discernible historical layers: those vowels inherited from Austronesian, which form the core of the basic vowel system, and

---

[2] However, it appears that this four-way Austronesian vowel distinction was already on its way to becoming a three-way distinction in parts of Western Malayo-Polynesian.

[3] It needs to be pointed out, however, that some of the expansion of the vowel inventory is due to borrowing from MK.

---

those vowels which primarily reflect MK influence and overwhelmingly occur in pre-Chamic MK borrowings.

While distinguishable, the two layers are not completely distinct: sometimes the phonology of the MK borrowings matched the phonology of the Austronesian lexicon, making the borrowed form indistinguishable on purely phonological grounds from inherited Austronesian forms; undoubtedly, sometimes the phonology of the MK borrowings was restructured by the pre-Chamic speakers to match the phonology of the Austronesian lexicon, again making the forms blend phonologically with the inherited Austronesian forms; but, in a way that is at times strikingly obvious, sometimes new phonological contrasts accompanied the MK borrowings.

Among forms carrying new phonological contrasts, the overwhelming majority of the words are identifiable as MK loans into pre-Chamic, while the bulk of the remaining forms are potentially of MK origin as they lack etymologies, Austronesian or otherwise. However, although overwhelming preponderance of forms containing new vowels are MK borrowings or possible MK borrowings, sprinkled in among the MK forms, there are also usually one or two words with straightforward, well-attested Austronesian etymologies. Two things appear to have happened in such words. First, the MK contact led to the development and phonemicisation of a vowel distinction already present in the phonetics of the Austronesian forms. Second, the development of the new sound in an Austronesian form would have significantly lessened the need to restructure the incoming MK loan words containing this vowel.

1.1 THE LITERATURE

The literature on Chamic vowels contains considerable discussion of the correspondences between An and PC (e.g. Blood 1962, Pittman 1959 and Thomas 1963), as well as a more limited discussion of the reflexes between PC and the modern Chamic languages (primarily Lee 1966, but also Burnham 1976 and others). However, two more recent developments make it possible to clarify, expand upon, and, in some cases, revise this earlier work. First, there has been a greatly expanded awareness of precisely which forms are MK borrowings; the use of Headley (1976) augmented by preliminary reconstructions of two branches of Mon-Khmer found in Vietnam (H. Blood 1968; Smith 1972) not only has allowed the recognition of a large number of MK loans but also—in conjunction with other revisions— has made it possible to work out a rough chronology for many of the loans, classifying them as either pre-Chamic or post-PC loans. Second, the database has expanded enormously, leading to numerous revisions in the individual lexical items and some modification in the overall schema, although much of Lee's outline is still quite workable today. In addition, of course, other recent literature in Chamic studies has also contributed to our understanding of PC vowels and their correspondences (e.g., Ni 1988a, 1988b, 1990a, 1990b; Haudricourt 1984; Benedict 1984; Blust 1969, 1980a, 1980b, 1981, 1983–84, 1986, 1989).

The only work to set out main vowel correspondences of PC was Lee (1966). Since then expanded knowledge of various Chamic languages makes the time appropriate for revisions. One source of revision is the realisation that some 10% of Lee's reconstructed forms are MK borrowings, many of them not even pre-Chamic borrowings but post-PC borrowings and thus are not legitimate input to PC reconstruction. The removal of these post-PC borrowings from the database eliminates certain of Lee's marginally attested vowel correspondence

patterns completely while simplifying others. A second development leading to the modification of Lee's protoforms is a reanalysis of his treatment of nasalised vowels. Lee often reconstructed nasalised vowels to account for the failure of certain Roglai word-final nasals to denasalise. However, the elimination of late borrowings from the database allows an alternative treatment of the Roglai patterns, which in turn makes it clear that the Roglai changes are internal to Roglai and, thus, no longer reconstructable to PC. Finally, as the result of modifications in the treatment of numerous individual words, it has been possible to reconstruct *ɔ where Lee reconstructs both *ɔ and *o. This treatment reassigns the various *ɔ reflexes largely to *ə, but occasionally elsewhere. As a general consequence of the accumulation of numerous minor revisions, this paper provides new reconstructions of the PC main vowel reconstructions, modifying Lee (1966).

As has already become obvious, this discussion of PC main-syllable vowels divides the relevant discussion into four time periods: the Austronesian period predating contact with MK languages; the pre-Chamic period, in which early contact occurred but which predates what we reconstruct as PC; the stretch of time during which what we reconstruct as PC Chamic was spoken; and the period following the break-up of PC.

1.2 PC VOWEL LENGTH

As will become clear later, vowel length in PC involves the interaction of the Austronesian inherited vowels with the MK vocalic contributions to PC. Here it is enough to make several comments on the distribution of vowel-length contrasts. In PC, vowel length occurs only for specified vowels and then only in certain contexts. As Lee (1966:117) noted, the "length contrast seems to be fairly certain for *a, *u, and *ɔ, but (as is true of the daughter languages) is limited to certain environments". The questions in the reconstruction of length revolve around determining precisely those finals before which length occurs and those before which it does not occur. The PC *a occurs both long and short before final -ʔ, -ŋ, -k, -l, -r, and marginally before -t (see Table 26). The PC *ɔ occurs both long and short before final -ʔ, -ŋ, and -k (see Tables 36–40). The PC *u occurs both long and short before final -ʔ and -ŋ (see Tables 11 and 13–16).

Other residual evidence of vowel length seems to exist in various daughter languages but it is not (yet?) possible to reconstruct it. For instance, the length distinctions in Rade suggest that there may have once been a distinction between -aːm and -am; however, if so, it has been totally obscured elsewhere by subsequent developments throughout Chamic.

TABLE 1: CONSTRAINTS ON THE OCCURRENCE OF PC VOWEL LENGTH

*-ă- versus *-aː-   /___ʔ; /___ŋ; /___k; /___l; /___r; /___n; /___t (marginal)

*-ɔ̆- versus *-ɔː-   /___ʔ; /___ŋ; /___k

*-ŭ- versus *-uː-   /___ʔ; /___ŋ

Notes on tables:

a) An in these tables refers to an Austronesian reconstruction that at least predates Chamic; many of these forms, of course, do not reconstruct all the way back to Austronesian. Two levels of borrowed entities are distinguished: borrowings

predating PC are marked by ^ǂ* with the ^ǂ indicating that the form was borrowed and the * indicating that nonetheless it reconstructs back to PC. Borrowings postdating PC are simply marked by ^ǂ. Most likely all the *ə forms should be prefaced in one of these ways.

b)   Apparent irregularities in the correspondences are indicated by a hyphen followed by a consonant indicating precisely what is irregular: -v = irregular vowel, -c = irregular consonant, -f = irregular final, -vr = irregular vowel register, -t = irregular tone, -ⁿ = irregular nasalisation, -l = irregular length, -iv = irregular initial and vowel, -ivf = irregular initial, vowel, and final, -r = irregular correspondence for /r/, -vʄ = the initial vowel is irregular, and so on.

c)   The symbol (m) indicates metathesis.

d)   'Bahnar (AC)' refers to the Bahnar forms cited in Aymonier and Cabaton (1906).

2. THE PC MAIN-SYLLABLE VOWELS INHERITED FROM AUSTRONESIAN

The pre-contact Austronesian language that was to become Chamic had a vowel system consisting of four main vowels, occurring in either syllable, and three diphthongs, occurring only in the second syllable (see Table 2).

TABLE 2: An MAIN VOWEL REFLEXES IN PC

An second-syllable vowels		PC main-syllable vowels	
*i	*u	*-i-	*-u-
		*-əi	*-əu
*e [ə]	→	*ă	
*a		*a	
*ay *uy		*-ay	*-ui
*aw		*-au	

The reflexes of these Austronesian vowels in PC are straightforward for the most part, with the subsequent PC reflexes set out in the tables below. In certain cases, particular developments are discussed in more detail. The essence of the An > PC changes, however, is relatively simple. The two high Austronesian vowels underwent splits, diphthongising in final position, but remaining -i- and -u- in closed syllables; these developments are also further conditioned in minor ways by an apparent interaction with stress placement (see discussion at §2.1 below). Austronesian shwa became *ă before certain finals but merged with *a before others; this led to a length distinction between PC short *ă and PC long *a before the finals where *ă was maintained (see Table 3 and further discussion in §2.5).

The original An shwa is realised as PC short *ă (Table 3); note that the PC words reconstructed with shwa are not inherited from An, but instead are borrowed from MK! The realisation of An shwa as PC *ă, by introducing a contrast with PC *a, introduced a vowel-length distinction into PC.

TABLE 3: An *e [= ə] > PC *-ă- REFLEXES

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham	
*ajeng	*ă-	hədăŋ	hədăŋ	hadak	hədaŋ	---	t̪ăŋ	hat̪ăŋ; t̪aŋ	char-coal
taŋek	*ă-	kənă?	tənă?	tană?	tənă?	cənă?	tanə?	tanĭ?	cook
lemak	*ă-	emă?	rəma?-l	lumă?	ləma?	ləmĭa?	lamə?	limĭ?; lamĭ?	fat; grease
*le(m)-beng	*ă-	ɓăŋ	ɓăŋ	ɓak	ɓaŋ	ɓăŋ	---	băŋ	hole; door
*gatel	*ă-	kətăl	kətal	katan	kətal	kətăl	katăl	katăl	itchy
hiket	*ă-	---	?akă?; kă?	ika?	aka?	?akă?	kăk	ikă?; kă?	to tie
qulej	*ă-	hluăt (m)	hlăt; hluăt (m)	hula?	həla?	---	hlă?	hală?	worm; caterpillar

Note: Forms in the Austronesian column without an asterisk are from Blust; however, such forms are only claimed to predate PC. Austronesian forms with an asterisk are from a myriad of other sources.

The new length distinction occurred before the final consonants *-?, -ŋ, -k, -l, -r, and, more marginally, before -t, causing a distinction between short PC *ă (< largely from An *e [ə]) and long PC *a (< largely from An *a). In other environments, An *e [= *ə] merged at some point with PC *a, although with further research it still may be possible to extend the reconstruction of the PC vowel-length difference to additional environments. For example, the reflexes of An *-em in PC are almost always *-ăm, but nonetheless it has not been possible as yet to reconstruct a distinction between *-am and *-ăm. Perhaps later research will allow a vowel-length distinction to be teased out in this context, but this has not been done yet.

The next historical stage involves the break-up of PC into its daughter languages. The various PC vowel reflexes are relatively clear, making it possible to represent the changes fairly adequately in various tables (cf. Table 4). There are, of course, little oddities such as sporadic metathesis scattered throughout Chamic and instances here and there of unaccounted-for nasalisation (neither of which will be discussed here), but although interesting in themselves, these oddities are a very minor part of the vowel reflex patterns.

TABLE 4: REFLEXES OF PC INHERITED MAIN-SYLLABLE VOWELS

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
*i	*-əi	-ɛi	-əi	-əi	-ai	-əi	-ŏi; -ɨi	-ay	-ĕy
	*i-	-i-	-i-	-i-	-i-	-i-	-e-; -ei?; -i-	-i-	-i-
*u	*-əu	-ău	-əu	-əu	-au; -(i)ə	-əu	-ău; -ɨau	-au	-ŏw

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
	*u-	-u-	-u-	-u-	-u-	-u-	-o-; -ou-; -u-	-u-	-u-
*e	*a-	-a	-a	-a	-a	-a	-a; -ia	-a	-a
*uy	*ui-	-ui	-ui	-ui	-ui	-ui	-oi; -ui	-ui	-uy
*ay-	*ay-	-ie	-ai	-ai	-a:i[242]	-ai	-ai; -iai	-ai	-ay
*aw	*au-	-au	-au	-au	-au	-au	-a:u; -iau	-au	-aw
*a-	*a	a	a	a	a	a	a; ia	a	a

Note: Vowel length is not shown in this table and will be discussed later. The former existence of vowel registers in Haroi accounts for the dual Haroi reflexes for each PC vowel.

In this paper, only two of the more notable PC vowel reflex patterns are given further discussion. In Western Cham and Phan Rang Cham, PC *a and *ă have an interesting set of conditioned reflexes, which are discussed below in some detail (§2.5).

The other PC vowel reflexes of particular interest are found in Haroi, which has what Huffman (1976) termed 'restructured register'. These Haroi changes will not be discussed in any detail here, but they have been discussed elsewhere in the literature by others (Lee 1977; Burnham 1976) and by myself (Thurgood 1996, 1997). These fascinating vowel splits were the focus of some early work by Lee (1977) and by Burnham (1976), who both correctly deduced that the vowel splits correlated with the earlier presence of vowel registers. Some of the details of the conditioning factors still remain to be figured out.

2.1 REFLEXES OF PC *-i- AND *-i

In the stage from An to PC, the reflexes of the Austronesian high vowel *i split: in open stressed syllables, it became PC *-əi (the PC reflexes of which are reflected in Table 5), while in closed syllables (and, apparently, in unstressed open syllables), it remained *-i- (the PC reflexes of which are reflected in Table 6).

TABLE 5:   REFLEXES OF PC *-əi < An *-i

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
-i	*əi	-εi; -uε (m)	-əi	-əi; -uəi (m)	-ai	-əi	-ïi; -ɔ̌i	-ay	-ɛ̌y

[m = metathesis]

The fact the split was conditioned both by the openness of the syllable and by the presence of stress becomes much more obvious when the data in Table 7 is examined.

TABLE 6:   REFLEXES OF PC *-i- < An *-i-

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham
---	*i-	-ih	-ih	-ih	-ih	-ih; -ĭh; -ɪh; -ĕh; -ɛh	-ih	-ĭh; -ih

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham
---	*i-	-ĭm	-ĭm; -im	-ip	-im	-ĭm; -ɛm	-ĭm	-ĭm
---	*i-	-ĭp	-ip	-iuʔ	-iuʔ	-ip	-iuʔ	-ĭwʔ
---	*i-	-ĭʔ	-ĭʔ	-iːʔ	-iːʔ	-iʔ; -eiʔ	-ĭʔ	-iʔ
---	*i-	-ĭt	-ĭʔ; -it	-iːʔ	-iːʔ	-iʔ; -ɛiʔ	-ĭʔ	-iʔ
---	*i-	-ĭl	-ĭl	-in	-il	-ĭl; -el	-ĭl	-ĭl
---	*i-	-ir	-ir	-i	-i	-ei	-i	-i
---	*i-	-ĭn	-in; -ĭn	-in; -it	-in	-ĭn; -ĕn	-ĭn	-ĭn
---	*i-	-ĭŋ	-ĭŋ	-iŋ	-iːŋ	-iŋ	-əŋ	-ĭŋ

The Austronesian open syllable *-i reflexes do not unexceptionally go to PC *əi. In a handful of 'grammatical' morphemes, the open-syllable *i displays a unique pattern of reflexes (see Table 7), albeit a pattern that matches the reflex pattern for *-i except for the Rade and the Haroi reflexes. In effect, except for the split Rade reflexes and the Haroi reflexes, the reflex pattern is the pattern for closed syllable *-i-.

TABLE 7: PC OPEN SYLLABLES WITH *-i IN UNSTRESSED SYLLABLES

The pattern:

Malay	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham
---	*-i	-ɛi	-i	-i	-i	-i; -ei	-i	-i

The examples:

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham	
di	*-i	ti	---	---	tə-	---	---	ţi	at
sini	*-i	tinɛi	---	tinĩ	ni	ni	ni	---	here
ini; ni	*-i	tinɛi	ʔanai	kunĩ	ni	ʔani	ni; nu ??	ni	this
---	*-i	---	---	---	---	tiʔ-v	---	---	particle
kami	*-i	həmɛi	gəməi	labuʔ	---	kəmei;	---	kami	we (exc.)
			-iv	kamĩn?		kəme -v			

The pattern immediately in Table 7 is quite exceptional, as the overwhelmingly dominant pattern for the word final *-i is for it to become *-əi in PC. Although there are several borrowed forms with similar reflexes, the forms in Table 7 are inherited forms, not borrowings. The set itself consists of several demonstratives, a particle, and a pronoun. A careful look at the syntax and semantics of these forms shows they all have something in common: they are all 'grammatical' forms and, more crucially, they are all typically unstressed, not stressed. Thus, the lack of stress in these forms seems to be the reason that these words have not patterned with the stressed An *-i > PC *-əi forms.

Actually, a more careful study of the variation in the Rade and the Haroi reflexes suggests, not that these forms were unstressed in every context, but rather that in some contexts the forms were stressed and in some they were unstressed, with either the stressed form or the unstressed form ultimately winning out on a case-by-case basis.

A small number of PC *-i* finals came from other sources. Two forms appear inherited from Austronesian (see Table 8).

TABLE 8: TWO OTHER INHERITED FORMS WITH PC OPEN SYLLABLE *-i*

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham	
sisir	*-i*	kəsi	təsi	kasi	təsi	cəsei	tasi	tathi	a comb
pagi	*-i*	məgi	pəgi	pagi	pəgi	pəkhι	pake	pake	tomorrow
		-v					-v	-v	

Headley (1976) identifies 'comb' as a borrowing from MK, but, if it is, its presence in Malay as *sisir* suggests that if it was borrowed, it was borrowed into An before PC; thus, the reflex still needs to be explained. In this case, the explanation lies in the final *-ir*. Paralleling Austronesian forms involving final *-ur* (§2.2), the change from An *-i* > PC *-əi* occurred before the change An *-ir* > PC *-i*; as a consequence, the *-i* did not undergo the change from *-i* > *-əi*.

However, with the form *pagi*, which shows up in Malay as *pagi*, there is no explanation for the unexpected PC final *-i*.

The remaining forms with a PC final *-i* all appear to be loans (Table 9). In some cases, Headley has identified it as a loan; in other cases, various other irregularities suggest it is a loan.

TABLE 9: APPARENT BORROWINGS WITH PC OPEN SYLLABLE *-i*

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham	
---	*-i*	ei	rəʔi	laʔi	---	lə--i	laʔi	lii	basket, winnowing
		(Headley (#1.5) identifies this as a MK loan.)							
---	*-i*	εa	həni	ia	həni	həni	ea	hani	bee;
		hənue (m)		hunī			hani		honey
		(cf. Vietnamese *ong mật*)							
---	*-i*	---	kli-i	---	---	cəlei	---	tali	flat (of large rocks)
---	*-i*	ki	təki	tuki	təki	cəke	take	take	horn; antler
		-v					-v	-v	
		(Headley (#1.34) identifies this as a MK loan; it also looks like a PLB form.)							
---	*-i*	kətrεi	---	katri	kətrəi	kətrŏi	katray	---	scissors
					-f	-vf			
		(Headley (#1.5) identifies this as a MK loan. Another source suggests *kĕlĕkati* 'areca-nut scissors' from Tamil.)							

Little question exists whether most, if not all, the above forms are loans. However, with at least several of the forms, there is some question about the relative chronology of the loan. Certainly, 'horn; antler' and 'scissors' look like loans that postdate the break-up of PC.

2.2 REFLEXES OF PC *-u, *-ŭ-, AND *-uː-

In the stage from An to PC, the reflexes of the An high vowel *u split: in open stressed syllables, it became PC *-əu (the PC reflexes of which are reflected in Table 10), while in closed syllables, it remained *-u- (the PC reflexes of which are reflected in Table 11).

TABLE 10: REFLEXES OF PC *-əu < An *-u

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
-u	*-əu	-ău	-əu	-əu	-au; -(i)ə	-əu	-ău; -ɨau	-au	-ɔ̆w

In closed syllables, An *-u- become PC *-u- (the PC reflexes of which are reflected in Table 11).

TABLE 11: REFLEXES OF PC *-u- < An *-u-

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
---	*u-	-uh	-uh	-uh	-u⁵⁵	-uh	-ŭh; -uh; -ŏh; -oh	-uh	-ŭh
-us	*u-	-uih	-uih	-uh	-u⁵⁵	-uːh	-ih; -oh	-uh	-ŭh
---	*u-	-ŭn	-ŭn	-ut; -un	-un	-un	-ŭn; -ŏn	-ŭn	-ŭn
---	*u-	-ŭm	-um; -ŭm	-up	---	-um	-ŭm; -ŏm	-um	-ŭm
---	*ŭ-	-ŭŋ	-uŋ; -ŭŋ	-uk; -uŋ	-uŋ	-uŋ	-ŭŋ; -ŏŋ; -oŋ	-uŋ	-ŭŋ
---	*uː-	-uŋ	-oŋ; -uŋ	-uːk	-uŋ	-ɔːŋ; -oːŋ	-uŋ; -oŋ; -oːŋ	-oŋ	-oŋ
---	*ŭ-	-ŭʔ	-ŭʔ	-uʔ	---	-uʔ	-ŭʔ; -ŏʔ; -ɔ̆ʔ	-ŭʔ	-ŭʔ
---	*uː-	-ŭʔ	-ŭʔ	-uʔ	---	-uːʔ	-uʔ; -ʊʔ; -ɔ̆uʔ; -oʔ	-uʔ; -oʔ	-uʔ; -oʔ

In addition to forms borrowed from MK, a small number of PC *-u finals are inherited from Austronesian (see Table 12).

TABLE 12: PC OPEN SYLLABLES FROM An *-ur FINALS

The pattern:

An	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham
*-ur	*u	-u	-u	-u	-u	-oːu	-u	-u

The examples:

Malay	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham	
ekor	*u-	ku	ʔaku	iku	aku	ʔakoːu	hla ku	---	tail
nyor	*u-	---	---	laʔu	ləʔu	ləʔu	laʔu	liu	coconut palm
kujur	*u-	kju	təju -i	---	---	kəsu	---	---	spear; lance

Malay	PC	Rade	Jarai	Roglai	Chru	Haroi	WCham	PR Cham	
telur (egg)	**u-*	---	---	---	*klu*	---	---	*klu*	scrotum (animal)

The pattern immediately above is interesting, as it reconstructs as PC **-u*, rather than the expected PC **-əu*. Quite obviously, in pre-Chamic word-final **-u* became **-əu*, and then the loss of final **-r* in **-ur* rhymes produced a set of new inherited word-final **-u* finals. In addition, of course, there are a number of other word-final PC **-u* forms, but probably these are mostly early loans.

It is significant that the developments in Acehnese closely parallel the developments in mainland Chamic languages. That is, although the normal Acehnese reflex for word-final **-u* is *-èe*, the word for 'tail' is *iku*, suggesting that the Acehnese forms paralleled the mainland Chamic forms in development. Although the evidence will not be given in this paper, work in progress substantiates the long-held belief that Acehnese is simply another Chamic language, albeit one that left the mainland at an early date.

In addition to the reflexes already discussed, **u* also has a limited vowel length distinction: **u* is found both long and short before final *-ʔ* (Tables 13 and 14) and final *-ŋ* (Tables 15 and 16). What is clear, however, is that at least some of the forms containing both the long and the short vowels are from Austronesian sources; it is also equally obvious that some of the forms in both categories come from MK sources.

One might suggest that, among the Austronesian descended forms, the short forms descend from **-uk*, while the long ones descend from **-ut*, as the single form in Rade, *mənŭt* 'banyan' with its final *-t* suggests, but without further data this is of course speculation. Similarly, the form *ribut* 'storm' ends in *-t* in Malay. However, at present, all this is essentially nothing more than unsubstantiated speculation.

TABLE 13: SHORT *-u-* BEFORE FINAL *-ʔ*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ŭʔ	*-ŭʔ*	*-ŭʔ*	*-uʔ*	*-uʔ*	*-ŭʔ*	
*ʔañŭʔ	*añŭʔ*	*ʔañuʔ* -vl	*añŭʔ*	---	*ñuʔ*	beads
*ñŭʔ	*ñŭʔ*	*ñŭʔ*	*ñŭʔ*	*ñuʔia*	*ñŭk* -f	dive; submerge
*mabŭʔ	---	---	*babuʔ*	*məbuʔ*	---	drunk
*manŭʔ	*mənŭʔ*	*mənŭʔ*	*manŭʔ*	*mənŭʔ*-ⁿ	*minŭʔ*	fowl; chicken
*ʔadŭʔ	*adŭʔ*	*ʔadŭʔ*	*aduʔ*	*aduʔ*	*aṭŭʔ; ṭŭʔ*	room
*bitŭʔ	*mətŭʔ*	*pətuʔ*	*pituʔ*	*pətuʔ* -f	*pitŭʔ; patŭʔ; patŭʔ*	star

TABLE 14: LONG -*u*- BEFORE FINAL -*ʔ*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-*uːʔ*	-*ŭʔ*	-*ŭʔ*	-*uːʔ*	-*uːʔ*	-*uʔ*; -*oʔ*	
*kaʔuːʔ	---	*kəʔŭʔ*	---	---	---	barkcloth
*guːʔ	*gŭʔ*	---	---	---	*ḳuʔ*	below, lower
*juːʔ	*jŭʔ*	*jŭʔ*	*juːʔ*	*juːʔ*	*çuʔ*	black
*yuːʔ	---	*yŭʔ* 'west'	---	---	---	descend
*taguːʔ	*kəgŭʔ*	*təguʔ*	*taguːk*-f	*təguːʔ*	*takoʔ*	get up
*tuːʔ	*tŭʔə*	*tŭʔ*	*tuːʔ*	---	*toʔ*	to receive
*ribuːʔ	*ebŭʔ*	*rəbŭʔ*	*rubuːʔ*	*rəbuːʔ*	*ripuʔ*; *rapuʔ*	storm
*kaʔuːʔ	---	*kəʔŭʔ*	---	*kəʔūːʔ-*[n]	---	worried; sad
*puːʔ	*pŭʔ*	---	*puːʔ*	*poːʔ*	---	carry in arms

Among the short vowels, there are several words with good Austronesian etymologies (apparently, for example, 'flour', 'mortar', and 'nose') as well as established MK borrowings. In contrast, at least preliminarily all the long vowels appear to be restricted to MK borrowings.

TABLE 15: SHORT -*u*- BEFORE FINAL -*ŋ*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
---	-*ŭŋ*	-*ŭŋ*; -*uŋ*	-*uk*; -*uŋ*	-*uŋ*	-*ŭŋ*; -*uŋ*	
*rabŭŋ	*ebŭŋ*	*reɓŭŋ* -i	*rubuk*	*rəbuŋ*	*ripuŋ*; *rapŭŋ*	bamboo shoot
*bŭŋ	*bŭŋ*	*buŋ*	---	*buŋ*	*puŋ*	basket, large
*tapŭŋ	*kəpŭŋ*	*təpŭŋ*	*tupuk*	*təpuŋ*	*tapŭŋ*	flour
*risŭŋ	*esŭŋ*	*rəsŭŋ*	*risuk*	*ləsuŋ* -i	*lithŭŋ*	mortar
*ʔidŭŋ	*adŭŋ*	*ʔadŭŋ*	*idūk*	*aduŋ*	*iṭŭŋ*	nose
*ʔanŭŋ	*anŭŋ*	*ʔanuŋ*	---	---	*anŭŋ*	package
*salŭŋ	---	*hlŭŋ*	*saluk*	---	---	pit, trench
*kadŭŋ	---	*kəduŋ*	---	*kəduŋ*	---	pocket; bag
*katŭŋ	*kətŭŋ*	*kətuŋ*	*katuk*	*kətuŋ*	---	pull
*tŭŋ	---	---	*tuk*	*tuŋ*	*tŭŋ*	stomach
*ɗŭŋ	*ɗŭŋ*	---	*ɗuk*	---	---	wrap up
⁺*phŭŋ*	*phŭŋ*	*phuŋ*	*phuŋ* -f	---	---	leper; leprosy
⁺ **gulŭŋ*	---	*gluŋ*	*paguluk*; *taguluk*	*pərləŋ*-v	*ḳalŭŋ*	to roll

(probably borrowed from Malay; see p.84 of Shorto)

⁺ *(li)hŭŋ*	*tɛi huŋ*-l	*huŋ*	*lahɔŋ* -f	*ləhɔŋ* - v	---	papaya

(Mon-Khmer; Headley, #1.46)

TABLE 16: LONG -u- BEFORE FINAL -ŋ

PC	Rade	Jarai	Roglai	Chru	PR Cham	
---	-uŋ	-oŋ; -uŋ	-u:k	-ɔ: ŋ; -o:ŋ	-oŋ	
*ʔamu:ŋ ?	amuŋ	---	amu:k	amu:ŋ- v; tərmung	---	bunch; ear; stalk
*caku:ŋ	kəkuŋ	cəkuŋ	caku:k	səko:ŋ	cakoŋ;takoŋ	carry
*ɗu:ŋ	ɗuŋ	---	ɗu:k	ɗo:ŋ	---	float
*ʔu:ŋ	uŋ	ʔoŋ	---	---	---	male; husband
*ʔamu:ŋ	amuŋ	---	mū:k -f	mɔ:ŋ	---	snout
*pu:ŋ	---	---	---	apo:ŋ	---	straw (rice)
*bru:ŋ	bruŋ	broŋ	---	---	---	streaked; striped

All the examples of long -u- before -ŋ appear to be from MK, but this remains speculation until more work has been done.

2.3 REFLEXES OF PC *-ay AND *-au

The reflexes of PC *-ay and *-au are straightforward and well-attested. The only particularly interesting reflex is in Tsat, where the final *-y strengthens to a glottal stop.

TABLE 17: REFLEXES OF PC*-ay AND *-au

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
-ay	*ay	-ie; -ai (loans)	-ai	-ai	-a:ʔ; -a:i (loans)	-ai	-ai; -iai	-ai	-ay
---	*au	-au	-au	-au	-a:u; -au	-au; -a:u	-a:u; -iau	-au	-aw
						(length is with apparent loans)			

The other interesting dimension to the reflexes of these two vowels is that they are, in effect, the long counterparts to the word-final PC *-əi and *-əu, which developed word-finally from An *-i and *-u, respectively.

2.4 REFLEXES OF PC *-ui

The reflexes of PC *-ui are straightforward.

TABLE 18: REFLEXES OF PC *-ui

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
-uy	*ui-	-ui	-ui	-ui	-ui	-ui	-ui; -oi	-ui	-uy

2.5 REFLEXES OF PC *a, *-ă-, AND *-aː-

Although sometimes subject to minor variation conditioned by the syllable-final consonant, the reflexes of PC *a in open syllables, and *-ă- and *-aː- in closed syllables are quite regular in PC.

TABLE 19: REFLEXES OF PC *-a, *ă, and *aː

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
---	*-a	-a	-a	-a	-a	-a	-ia; -a	-a	-a
-ah	*a-	-ah	-ah	-ah	-a[55]	-ah	-ah; -iah	-ah; -ih	-ah; -ih
-as	*a-	-aih	-aih	-a	-a	-aːh	-ah; -iah	-ah	-ăh
---	*a-	-am; -ăm	-am; -ăm	-ap; -am	---	-am	-iam; -ăm; -am	-ăm	-ăm
---	*a-	-ăp	-ăp	-aʔ; -ăʔ	---	-aʔ; -āuʔ	-auʔ	-auʔ	-ăʔ
---	*ă-	-ăŋ	-ăŋ	-ak	---	-aŋ	-ăŋ	-ăŋ	-ăŋ
---	*aː-	-aŋ	-aŋ	-aːk	---	-aːŋ	-aːŋ; -aŋ	-aŋ; -iŋ	-aŋ; -ăŋ (occasional)
---	*ă-	-ăʔ	-ăʔ	-aʔ	---	-aʔ	-ăʔ; -iaʔ	-ăʔ; -əʔ	-ăʔ; -iʔ
---	*aː-	-ăʔ	-ăʔ	-aːʔ	---	-aːʔ	-aʔ	-aʔ	-aʔ
---	*ă-	-ăk; -ăʔ	-ăk; -ăʔ	-ak; -aʔ	---	-aʔ	-ăʔ	-aʔ	-aʔ; -ak
---	*aː-	-ak	-ak	-aːʔ	---	-aːʔ	-aʔ	-aʔ; -iʔ	-aʔ; -iʔ
---	*ă-	-ăr	-ăr	-a	---	-ar	-al; ---	-ăr	-ăr
---	*aː-	-ar	-ar	-a	---	-aːr	-al; ---	ar	-ăr; -ar
---	*ă-	-ăl	-al	-an	---	-al	-ăl	-ăl	-ăl
---	*aː-	-al	-al	-an	---	-aːl	-al	-al	-al
---	*ă-	-ăn	-ăn	-an; -at	---	-an	-ăn	-ăn; -in	-ăn
---	*aː-	-an	-an	-aːn; -aːt	-aːn	-aːn	-ian; -an	-an; -in; -in	-an; -in; -in
---	*ă-	-ăt	-ăt; -ăʔ	-aʔ	---	-aʔ	-ăʔ	-ăʔ	-ăʔ
---	*aː-	-at (marginal)	-at; -aʔ	-aːʔ	---	-aʔ	-at; -aʔ	-aʔ	-aʔ
---	*ac	-ač	-ăiʔ	-aːiʔ	-aiʔ	-aːiʔ	-aiʔ; -iaiʔ	-aiʔ	-ayʔ

The marginally attested pattern noted above refers to the contrast between long and short *-a- before a final *-t. More evidence may strengthen this correspondence, or, conversely, eliminate it.

The PC *a occurs both long and short before final -ʔ, -ŋ, -k, -l, -r, and marginally before -t (see Table 26). It is widely suggested in the literature that the length distinction in PC correlates with certain Acehnese vowel distinctions; a careful examination of the two fully substantiates that claim for these vowels.

TABLE 20: LONG AND SHORT -*a*- BEFORE FINAL -*ʔ*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ăʔ	-ăʔ	-ăʔ	-aʔ	-aʔ	-ăʔ; -ɨʔ	
*hunăʔ	---	---	hunāʔ	hənaʔ	hanɨʔ	asthma
ǂ *băʔ	ɓăʔ	băʔ	baʔ	baʔ	păʔ	carry on back (borrowing?)
*tăʔ	---	tăʔ	taʔ	taʔ	---	chop; cut
*tanăʔ	kənăʔ	tənăʔ	tanāʔ	tənaʔ	tanɨʔ	cook
ǂ *lamăʔ	emăʔ	rəmaʔ -l	lumāʔ	ləmaʔ	limɨʔ; lamɨʔ	fat, oil, grease

(Bahnaric *rəmā, ləmā*)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*paghăʔ	---	pəkhăʔ	pakə́ʔ-v	khaʔ	ḳhăʔ	forbid
*ŋăʔ	ŋăʔ	ŋăʔ	ŋāʔ	ŋaʔ	aŋăʔ cf. ŋăʔ	make, do
*ʔurăʔ	arăʔ	ʔarăʔ	---	araʔni	urăʔ; arăʔ	now
*tisăʔ	kəsăʔ	təsăʔ	tisaʔ	təsaʔ	tathăʔ	ripe; cooked
*tapăʔ	kəpăʔ	təpaʔ -l	tupaʔ	təpaʔ	tapăʔ	straight; honest
*măʔ	măʔ	măʔ	māʔ	maʔ	mɨʔ	take; get
*ʔikăʔ	---	ʔakăʔ; kăʔ	ikaʔ	akaʔ	ikăʔ; kăʔ	to tie
ǂ *surăʔ	hrăʔ	hrăʔ	suraʔ	sraʔ	harăʔ	write; book

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-a.ʔ	-ăʔ	-ăʔ	-a.ʔ	-a.ʔ	-aʔ	
*tana.ʔ	---	tənăʔ	---	---	---	faggot; bamboo strip
*pa.ʔ	păʔ	păʔ	pa.ʔ	pa.ʔ	paʔ	four
*ja.ʔ	djăʔ	ʔjăʔ	dja.ʔ	---	---	hold; carry
ǂ *ɓla.ʔ	---	---	ɓla.ʔ	ɓla.ʔ	ɓlaʔ	open eyes wide
*pala.ʔ	plăʔ	plăʔ	pala.ʔ	pla.ʔ-vl	palaʔ	palm; sole

TABLE 21: LONG AND SHORT -*a*- BEFORE FINAL -*ŋ*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ăŋ	-ăŋ	-ăŋ	-ak	-aŋ	-ăŋ	
*glăŋ	dlăŋ	---	---	---	ḳlăŋ	look at
*lawăŋ	---	---	---	rəwaŋ	rawăŋ -i	visit
*hadăŋ	hədăŋ	hədăŋ	hadak	hədaŋ	hatăŋ	charcoal
*khăŋ	khăŋ	khăŋ	khak	khaŋ	khăŋ	hard; stiff

(Vietnamese ?)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*ɓăŋ	ɓăŋ	ɓăŋ	ɓak	ɓaŋ	băŋ	hole; door
ǂ *hăŋ	hăŋ	hăŋ	hak	haŋ	hăŋ	hot; spicy

(Mon-Khmer; Headley, #1.35 & Shorto)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*wăŋ	wăŋ	---	wak	waŋ	wăŋ	sickle
*păŋ	---	păŋ	---	---	---	wall, make
*lanăŋ	enăŋ	rənăŋ	lanak -f	lənaŋ	lanĭŋ	earthworm
*-a:ŋ	-aŋ	-aŋ	-a:k	-a:ŋ	-aŋ; -ăŋ	
*plăŋ	plăŋ	---	plak	pla:ŋ -l	---	lemon grass
*dra:ŋ			dra:k	dra:ŋ		rhinoceros hornbill
*kra:ŋ	----	kra:k	kəra:ŋ -v¹		---	clam; shellfish
⁺raba:ŋ/n -if	kəban -f	---	raba:k	rəba:ŋ	ripaŋ; rapaŋ	bridge
(Mon-Khmer; Headley, #1.10)						
*pina:ŋ	mənaŋ	pənaŋ	pinãŋ -l	pəna:ŋ	paniŋ	betel; betel-nut
*cana:ŋ	---	cənaŋ	canãŋ	chəna:ŋ	taniŋ	bed
*da:ŋ	daŋ	daŋ	da:c	da:ŋ	daŋ	(lie) on back
*ha:ŋ	haŋ ɛa	haŋ	ha:k	---	---	bank; shore
⁺*kala:ŋ	tlaŋ	klaŋ	kala:k	kəla:ŋ	---	hawk; kite
(Mon-Khmer; Headley, #1.38)						
*tula:ŋ	klaŋ	təlaŋ	tula:k	təla:ŋ	talaŋ	bone
⁺*ka:ŋ	kaŋ	kaŋ	ka:k	təlka:ŋ; ka:ŋ	kaŋ	chin; jaw
(Mon-Khmer; Headley, #1.12)						
*cada:ŋ; *rada:ŋ	kədaŋ	cədaŋ	cada:k	chəda:ŋ; səda:ŋ	radăŋ -m	crack open
*pagha:ŋ	bhaŋ -i	pəkhaŋ	pakha:k			dry over fire
*rala:ŋ; *kala:ŋ	hlaŋ	hlaŋ	rala:k	rəla:ŋ	---	grass, thatch
⁺*ja:ŋ	---	---	ja:k	ja:ŋ	---	to guard; gate(way)
(Sanskrit; Coope)						
*luba:ŋ	---	---	luba:k	ləba:ŋ	lipaŋ; lapaŋ	hole; pit
*sa:ŋ	saŋ	saŋ	sa:k	sa:ŋ	thaŋ	house
*pisa:ŋ	---	---	pisa:k	pəsa:ŋ	pathaŋ	husband
(Sanskrit)						
*ʔura:ŋ	arăŋ -l	ʔarăŋ	ura:k	ara:ŋ	uraŋ	person
*huda:ŋ	hədaŋ	hədaŋ	huda:k	həda:ŋ	hatăŋ; taŋ	shrimp; lobster
*ya:ŋ	yaŋ	yaŋ	ya:k	ya:ŋ	yaŋ	spirit; god
*la:ŋ	laŋ	laŋ	la:k	la:ŋ	laŋ	spread out
*luwa:ŋ	ewaŋ	rəwaŋ	luwa:k	ləwa:ŋ	liwaŋ	thin; lean
*caɓa:ŋ	kəɓaŋ	---	jabha:k	chəɓa:ŋ	caɓaŋ	branch; fork
*kata:ŋ	kətaŋ	kətaŋ	---	kədaŋ -l	---	strong; well

PC	Rade	Jarai	Roglai	Chru	PR Cham	
ǂ *piɲan	məŋan	---	---	---	paɲĭn-	dish

(Hindi < Persian; Coope)

TABLE 22: LONG AND SHORT -a- BEFORE FINAL -k

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ăk	-ăk; -ăʔ	-ăk; -ăʔ	-ak; -aʔ	-aʔ	-aʔ; -ak	
*tagăk	kəgăʔ	təgăʔ	tagak	təgak -f	---	bush knife
*ʔalăk -f	---	---	alak	alak -f	alăk	wine; liquor
(Arabic)						
*taglăk	kədlăk	təglăk	---	tərglaʔ	---	choke
*gatăk	kətăk	kətăk	kataʔ	kətaʔ	katăʔ	sap; resin
*parăk	prăk	prăk	pariaʔ	priaʔ	paryăʔ;	silver;
					pirak	money
					(white)	
*ʔalăk	alăʔ	ʔalăʔ	---	---	---	yolk
*-a:k	-ak	-ak	-aːʔ	-aːʔ	-aʔ; -iʔ	
*rata:k	etak	rətaʔ	rataʔ	rətaʔ	riṭaʔ; rataʔ	bean; pea
*paʔa:k	păl-ak	pəʔăʔ	paʔaʔ	pəʔaʔ	paaʔ	armpit
*ʔana:k	anak	ʔanăʔ-l	anãᵐ	anaʔ	aniʔ	child
ǂ *ʔa:k	ak	ʔaʔ	aːʔ	chim aːʔ	aʔ	a crow
(Mon-Khmer)						
*maña:k	---	---	mañãⁿʔ	---	miñiʔ	oil
(from Cham? Shorto p.87)						
ǂ *ʔamra:k	amrak	---	amraʔ	amrãːʔ-ⁿ	mraʔ;	peacock
(Mon-Khmer; Headley, #1.47)					amraʔ	
*ha:k	hak	hak	haːʔ		split	
ǂ *ʔawa:k	awak	---	---	awaʔ	awaʔ	spoon; ladle
(Mon-Khmer; Headley, #1.61)						
*ja:k	jak	jak	jaːʔ	---	çaʔ	invite

TABLE 23: LONG AND SHORT -a- BEFORE FINAL -l

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ăl	-ăl	-al	-an	-al	-ăl	
*taŋăl	kəŋăl	---	---	---	taŋŏh -fv	deaf
*patăl	---	---	---	pətal	patăr	pillow
*gatăl	kətăl	kətal	katan	kətal	katăl	itchy
*sapăl	păl	həpal	sapan	spal	hapăl	arm (fore-)
*ɗăl	---	ɗăl	---	---	---	to wedge

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*măl	---	mal	mān	mal	măl	beam
*sanăl	anal-	hənal	sanān	---	---	pillow
*-aːl	-al	-al	-an	-aːl	-al	
*praːl -v	pral	prăn	pran	praːn	prɨn	strong; well
⸶ *jaːl	jal	jal	jan	jaːl	çăl -l	net, casting

(Indo-European; Headley, #2.6; Sanskrit *jaala*)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*kapaːl	kəpal	kəpal	kapan	kəpaːl	kapal	thick
*kaːl	kal	kal	---	kal -l	---	to lock; bolt
*baːl	---	bal	ban	baːl	pal	mend; patch
*banaːl	mənal	bənal	banān	---	---	rag
*kataːl	---	---	katan	kətaːl	katal	thunder; lightning

TABLE 24: LONG AND SHORT -*a*- BEFORE FINAL -*r*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ăr	-ăr	-ăr	-a	-ar	-ăr	
⸶ *ʔasăr	asăr	---	---	---	athăr	seed

(Sanskrit *saara*)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*padăr	---	pədăr; pədər	pada	pədar	---	spin; turn
*paɗăr	---	---	paɗa	pəɗar	---	tell, send
*-aːr	-ar	-ar	-a	-aːr	-ăr; -ar	
*baʔaːr	məar	---	baʔa	bəʔaːr	piar	paper
*ɓaːr	---	---	ɓa	---	ɓăr	coiled
*daːr	dar	dar	---	---	---	encircle
*ʔusaːr	---	ʔasar; ʔasăr	usa	asaːr	athăr	flesh, meat
*chaːr	čhar	---	cha	saːr	char	gong
*waːr	war	war	wa	waːr	wal -f	stable; pen

TABLE 25: LONG AND SHORT -*a*- BEFORE FINAL -*n*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-an	-ăn	-ăn	-an; -at	-an	-ăn	
*ʔapăn	---	---	apat	apan	apăn; păn	hold; take
*ŋăn	ŋăn	ŋăn	---	---	---	money
*ʔaŋăn	---	---	---	aŋan	aŋăn; ŋăn	name
*ʔanăn	anăn	ʔanăn	anān	---	---	name
⸶ *klăn	tlăn	klăn	tlat	klan	klăn	boa; python

(Sanskrit ?) (for *ʔaŋăn row)

(Mon-Khmer; Headley, #1.50)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
⚹ *khăn	---	khăn	khat	khan	khăn	cloth

(Vietnamese *khăn* 'towel; handkerchief')

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*găn	găn	găn	gat	---	ḳăn	cross; pass over; go past
*-a:n	-an	-an	-a:n; -a:t	-a:n	-an; -in; -in	
*lama:n	eman	rəman	lumān	ləma:n	limin	elephant
*taŋa:n	kəŋan	təŋan	taŋān	təŋa:n	taŋĭn	hand
⚹ *papa:n	---	---	---	pəpa:n	papan	board; plank

(probably borrowed)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*dha:n	adhan	than; dhan	tha:t	tha:n	ṭhan	branch
*lupa:n	epan	rəpan	lupa:t	ləpa:n	lipan; lapan	centipede
*ʔika:n	kan	ʔakan	ika:t	aka:n	ikan	fish
*riña:n	eñan	rəñan	riña:t	ləya:n	liñan; lañan	ladder
*bula:n	mlan	blan	ia bila:t	ɛa bla:n	pilan	moon; month
*_na:n	tɛinan-i	pənan	ina:t; rina:t	---	---	pineapple
*huja:n	həjan	həjan	huja:t	həja:n	haçan	rain
*jala:n	elan	jəlan	jala:t	jəla:n	çalan	road; path
*bha:n -ⁿ	---	phan	pha:t	phā:n -ⁿ	phan	sneeze

TABLE 26: LONG AND SHORT *-a-* BEFORE FINAL *-t*

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ăt	-ăt	-ăt; -ăʔ	-aʔ	-aʔ	-ăʔ	
*sukăt	kăt	həkŏt-v	sukaʔ	skaʔ	---	stopper
*pisăt	məsăt	---	pisaʔ	pəsaʔ	pathăʔ	navel; centre
*buŋăt	məŋăt	bəŋăʔ; bəŋăt	buŋāʔ	bəŋaʔ	piŋŭʔ	soul, spirit; shadow
*ʔurăt	aruăt (m)	ʔaraʔ -v	uraʔ	araʔ	urăʔ	vein, tendon
*hulăt	hluăt (m)	hlăt; hluăt (m)	hulaʔ	həlaʔ	halăʔ	worm
*kawăt; *kuat	kəwăt	kuăt	---	kuat -f	---	wire
⚹ talabăt	---	---	talabat -f	təlbat -f	---	worship

(borrowing?)

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-a:t	-at	-at; -aʔ	-aʔ	-aʔ	-aʔ	
*pha:t	məñat	phăʔ	pañāʔ	phaʔ	phaʔ	chisel
*jaha:t	jhat	sat	---	jəha:ʔ; jəhua	çaʔ	bad; wicked

PC	Rade	Jarai	Roglai	Chru	PR Cham	
⁺la:t	*lat*	*lat; la?*	*la?*	---	*klet*	flat
(Mon-Khmer; Headley, #1.25)					-ivf	
*laba:t	*ebat*	---	*luba?*	---	*lipa?; lapa?*	walk, go

In Table 27, which shows PC forms with **a* following a nasal consonant, the corresponding Western Cham and Phan Rang vowels are /-ɨ-/ or, more rarely, /i/ as the result of the nasalisation. For three forms ('flower', 'striped' (which may be the same root as 'flower'), and 'soul, spirit; shadow'), the reflex is /-u-/, a reflex conditioned both by the word-final position after /ŋ/ and by the fact that the initial **b-* gives the whole word second register.

TABLE 27: NASALISED **a* IN W. CHAM AND PHAN RANG CHAM

PC	WCham	PR Cham	
*tiŋa	---	*tañi*	ask
*tə(ri)ŋa	*(tanɨh)*	*taŋi*	ear
*mamah	*mamɨh*	*mɨmɨh*	chew
*tanah	*tanɨh*	*tanɨ̆h*	earth, soil
*tama	*tamɨ*	*tamɨ*	enter
*?ama	*mɨ*	*amɨ*	father
*huma	*hamɨ*	---	field, dry
*lima	*lamɨ*	*limɨ;lamɨ*	five
*?ina	*nɨ*	*inɨ*	mother; major; big
⁺/*?amāh	*mɨh*	*mɨ̆h*	gold
⁺lamāh	*ramɨh*	*ramɨ̆h* -i	rhinoceros
*laŋa -v	*laŋi*	*liŋi;laŋi*	sesame
*panah	*panɨh*	*panɨh*	shoot (bow); a bow
*mañam -l	*mañim*	*mɨñim*	weave; twill
*pina:ŋ	*panɨŋ*	*panɨŋ*	betel (areca palm); betel nut
*?ana:k	*nɨ?*	*anɨ?*	child
*lama:n	*lamin*	*limin*	elephant
*cana:ŋ	*canɨŋ* (wooden bed)	*tanɨŋ*	furniture; bed
*taŋa:n	*taŋən* -v	*taŋɨ̆n*	hand
*maña:k	*mañi?*	*mɨñi?*	oil
*makrāh	*krɨh*	---	middle, half
⁺krām -lvⁿ	*krɨm* -v	*krɨm* -v	bamboo
*krāh	*krɨh; kɨh*	*krɨh*	middle; half
*gunam	---	*ḳanɨ̆m*	cloud
*nam	*nɨ̆m*	*nɨ̆m*	six
*?anɨ̆n	*nən*	*nɨ̆n*	that (third p.)

PC	WCham	PR Cham	
*buŋa	paŋur -f	piŋu	flower
*buŋa	---	piŋu	striped
*buŋăt	paŋŭʔ	piŋŭʔ	shadow, shade; soul, spirit

The type of nasalisation is itself of interest, being perseverative rather than anticipatory, a type of nasalisation reminiscent of the nasalisation in Malay and in other Austronesian languages.

3. THE PC MAIN VOWELS BORROWED FROM MK

The main vowels reconstructed for PC, including early pre-Chamic MK borrowings reconstructable to the PC level, are presented in Table 28 below. The MK borrowings are in bold.

TABLE 28: MAIN VOWELS BORROWED FROM MK (IN BOLDFACE)

PC main-syllable vowels

*ia   *-i-, *-i   *-u   ***ua**
⁺iău   *-ŭ-,   *-uː-   ***uai**
*iau   ***uəi**
                          *-əi,   *-əu,
*ɛ   *ə   *-ɔ
                                          *-ŏ-,   *-ɔː-

PC main-syllable vowels

                              *-a
                  *-ă-,   *-aː-
                        *-ay   *-ui
                        *-au

In Table 28, the PC vowels that came with MK borrowings are represented in bold type. This statement, however, requires some qualification. First, the *-i* in open syllables originates from two sources, one involving borrowed MK forms, the other involving forms inherited from Austronesian. It is important to recall that forms with *-i* in a final open syllable inherited from An became *-əi* in PC. Thus, MK borrowings with *-i* in a final open syllable that postdate this change introduced a number of new forms with *-i* in final open syllables. Further, a small number of Austronesian 'grammatical' forms apparently did not undergo the change from *-i* to *-əi* in unstressed contexts (see discussion in §2.1). Second, the *ɛ* is quite marginally attested; further research may eliminate it completely. And, third, the vowel-length distinction with *-u-* seems to have come about through the influence of MK borrowings containing long *-u-*.

3.1 REFLEXES OF PC *e

There are only a small number of PC forms that reconstruct with *e and their etymological status is not completely clear.

TABLE 29: REFLEXES OF PC *e

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
---	*ɛ-	-ɛh	-ɛh	-eh	-e	-ɛh	-ɪh; -ĭh; -iĕh; -ɛh	-ɛh	-ɛh
---	*ɛ-	-ĕ	-e	-e	-e	-ɛ	-ɪ; -ɛ	-ɛ	-ɛ
---	*ɛ-	-al	---	-en	---	-ɛl; -ɛːl	---	---	---

The majority are borrowings, but there is one obvious non-borrowing among them, the form *labeh 'more, surplus', which is obviously related to Malay lebih 'more'.

3.2 REFLEXES OF PC shwa

PC shwa only occurs in borrowed forms; the An *e [ə] became, not shwa, but PC *ă.

TABLE 30: REFLEXES OF PC *ə

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
---	*ə-	-ɨʔ	-ɨʔ	-əʔ	-əʔ	-əʔ	-ɔ̆ʔ; -əʔ	-əʔ	-əʔ; -ɨʔ
---	*ə-	-ih	-əh; -ih	-əh	---	-əh	-ɨh;-ɔ̆h; -ɔ̆h	-əh	-ɔ̆h; -əh
---	*ə-	-ɔ̆k	-ɨk; -ɨʔ	-əʔ	-əʔ	-əʔ	-ɨʔ;-ɔ̆-	-əʔ	-ɔ̆ʔ
---	*ə-	-ăm	-ɔ̆m	-əp	-an	-əm	-a-; -ĭa-	-ɔ̆m	-ăm
---	*ə-	-ɔ̆ŋ	-ɔ̆ŋ	-ək	-aŋ	-əŋ	-ɨŋ;-iŋ	-əŋ	-aŋ; -ɨŋ
---	*ə-	-ăn	-ɔ̆n	-ət	---	-ən	-ɔ̆n; ǂ-ŭn	-ĕn	-ăn
---	*ə-	-ɔ̆ʔ	-ɔ̆ʔ	-əʔ	---	-əuʔ	---; ---	-auʔ	-ɔ̆wʔ
---	*ə-	-ər	-ər; -ɔ̆r	-ə	-an	-ər	-ʊl;-ɔl	-ăr	-ăr

All of the forms with the above vowel reflex patterns (Table 30) and with clear etymologies (Table 31) are borrowings. There are, of course, numerous forms without clear etymologies.

TABLE 31: SOURCES OF PC *ə

PC	P-Mnong	P-North Bahnaric	Bahnaric (AC)	
ǂnrən -if; ǂdrə̄n -if	---	---	---	numb
ǂlaʔən	*ʔlik	---	---	cold

PC	P-Mnong	P-North Bahnaric	Bahnaric (AC)	
*grəm -vf	---	---	grâm	thunder

(Mon-Khmer; Headley, #1.66)

*chər	---	---	---	plant with stick
*ch-an-ər	---	---	---	dibble stick
*bə?	---	---	---	to fence, dam
*b-an-ə?	---	---	bənot	a dam, fence
*pəŋ	*pə̆ŋ	---	---	to nail, hammer
*jəŋ -vf	*jeŋ	---	---	become
*yəh	*yəh	---	---	particle
⁺*kləp	*tə̆p	---	---	stab; poke
⁺*pər	*pă̆r	*pă̆r	apăr; pă̆r	to fly

(Mon-Khmer; Headley, #1.27; Vietnamese *bay* from *bal*)

*ɓrəm	*kă̆m	---	brəm; mrəm	arrow
*gər	---	---	gər	handle (knife)
*gəŋ	---	---	ga:ŋ; gə̆ŋ	pole; post
⁺*sagər	---	*hagàr	həgər; çər	drum

(Mon-Khmer; Headley, #1.22)

*sidəm	---	---	huduṃ; hadaṃ	ant
*təl	---	---	tâl; tol	arrive; until

Although all the above forms are borrowings, it is not clear were all of them have come from. As the table makes clear, some have etymological connections to the Mnong branch of MK or to the Bahnaric branch. The first three forms are post-PC borrowings into various Chamic languages, as noted by the use of the symbol ⁺, but only one has a clear MK etymology. The next four forms also appear to be MK borrowings, as evidenced by the MK instrumental infix -an-; although it is possible to factor out this prefix on the basis of the forms in PC, it is far more likely that the forms were borrowed with the infixes already in place. The next eleven forms all have counterparts in Proto Mnong (Blood 1968), Proto North-Bahnaric (Smith 1972), or in Bahnaric itself (Aymonier and Cabaton 1906). None of the PC *ə forms appear to have Austronesian etymologies.[4]

3.3 REFLEXES OF PC *ia, *iāu, AND ⁺iau

Three diphthongs have been borrowed from MK into PC: *ia, ⁺iāu, and *iau. The reflexes are conditioned by co-occurrence with different finals, but are nonetheless quite regular.

---

[4]   Most likely all the *ə forms should be prefaced with the symbol ⁺, indicating a borrowing, with those borrowed from MK into pre-Chamic being indicated by ⁺* and those borrowed after the break-up of PC being indicated simply with ⁺.

TABLE 32: REFLEXES OF PC *ia*

PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham	
*-ia	-ia	-ia	-ia	-ia	-ia	---; -ĕa; -eă	-ea	-ya	
*ia-	-iă?	-iă?	-ia?	-ia?	-ia?	---; -ĕa?	-i?; ? -ea?	-yă?	
*ia-	-ier; -ɛa	-ɛr; -ia	-ia	-ia	-ia; -iər; -iɛr	---; -ɛa; -ea	-ia; -ea	-iɛr; -ier	
*ia-	-ia-	---	---	---	-iã:u?	-cau?	-eau?	-yaw?	
*ia-	-iă?; -iet	-iă?; -ɛt	-ia?	-ia?	-ia?	-ɨa?; -ɛt -v	-i?; -ea?	-?; -yă?	
*ia-	-iăm; -iam -iam	-iap	-ia?	-ia:m	---;	---; -eam	---; -eam	-yam	
*ia-	-ieŋ	-ɛŋ	-iaŋ; -iak (-yak)	-iaŋ	-iaŋ	---; -eaŋ	-iaŋ	-ieŋ	
*ia-	-iă?	-iă?	-ia?	-ia?	-ia?	---; -ĕa?	-i? ?; -ea?	-yă?	
*?iar	ɛa	?ia	ia	?ia[33]	ia	?ɛa; ?ea	ea	ýa; ier	water (fresh)

Of all the words containing the diphthong -ia- only 'water' (last item in Table 32) appears to be a Austronesian word. The overwhelming majority of all the above forms are borrowings, although *chiyap* 'wing' (last item in Table 33) may, despite its initial, be an inherited form at the PC level.

TABLE 33: REFLEXES OF THE PC CONFIGURATION *iya-*

PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham	
*iya-	-ia-	-ia-	-ia-	--- /l___; -ia:-	-əya- -ɛa-	---; -aya- /r___	-ea-; /r___; -ya-	-aya	
*chiyap	siap	---	---	---	siãu? _n	cau? -v	cheau?	thyaw?	wing

This last 'vowel' is obviously not a unitary vowel, but rather a specific configuration that appears to behave uniquely. In some cases, this particular configuration has coalesced into *ia*.

TABLE 34: REFLEXES OF PC *iãu*

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
---	*iãu-	-iău; -au	-ɛu; -iɛau;	-iãu	-a:i?	-iãu	-iau; -ɛau;	-iu	-iw

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
			-iau				-eau		
---	⧧*iau-	-ieo	---	-iãu	-iau	-iau	-iau;	-iau	-iew
							-ɛau		

The above patterns occur only in borrowings. The first pattern ⧧iāu is even further restricted; it occurs only in words borrowed after the break-up of PC, as the ⧧ before the form indicates.

3.4 REFLEXES OF PC *ua AND *uai

The overwhelming majority of the words in PC containing the above vowels are borrowings, but there are at least two forms that are inherited: *buat 'to do' and *dua 'two', both identical to the forms in Malay (see Table 35).

TABLE 35: REFLEXES OF PC *ua AND *uai

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
--	*-ua	-ua	-ua	-ua	---	-ua	-ua; -ɔa	-ɔa	-wa
---	*ua-	-uah	-uah	-uah	-ua⁵⁵	-uah	-uah; -oah	-ɔah	-wăh
---	*ua-	-ăt; -uot	-uă?; -ɔ̆t	-uã?	-ua?	-ua?	---; -ŏa?	---	---
---	*ua-	-uom	-ɔm	-oːp	-uam	-ɔːm	-uɔm; -ɔm	-ɔm	-ɔm
---	*ua-	-uon	-ɔn	-uat; -uan	-uan	-uan; -uən	-uan; -ɔn	-uan	-ŏn
---	*ua-	-uă?	-uă?	-ua?	-ua?	-ua?; -əwa?	-uă?; -oa?	-ɔa?; -uɪ̆?	-wă?
---	⧧*ua	-uor; -ua	-ua	-uã	-ua	-ua	-oa	-ur; -ɔa	-u
---	⧧*ua-	-ul	-ul	-uan; -uən	---	-ual; -uəl	-ul	-ual	-ŏl
---	⧧*uac	-uĕ?	-uăɪ̆?	-ue?; -uai?	-oi?	-uai?	-oai?; -uɔi?	-uai?	-ɔy?
---	⧧*uəi-	-ui	-ui	-uəi	-oi	-uəi	-ui; ---	-uai	-oy
---	⧧*uc	-uč	-uɪ̆?; -uc	-uɪ̆?	---	-uɪ̆?	-uɪ̆?; -ŏɪ̆?	-uɪ̆?	-ŭy?
---	⧧*uai-	-ue	-uai	-uai	-uaːi	-uai	-uai; -oai	-uai	-oy

As with a number of the correspondences examined so far, it is sometimes quite difficult to distinguish between conditioned variation and irregularities due to borrowing.

3.5 REFLEXES OF PC *-ɔ, *-ɔ̆-, AND *-ɔː-

The majority of the *-ɔ, *-ɔ̆- and *-ɔː- vowels entered PC through borrowing, but there are nonetheless a minority that appear to have come not through borrowing but from Austronesian forms with *u. In particular, *ramoːŋ 'tiger', *lamɔ or *ramɔ 'cow', *trɔ̆ŋ 'eggplant', and *dɔːk 'sit; stay; live' seem to have some claim to some sort of pre-Chamic Austronesian etymology; the etymology for *dɔːk is particularly strong.

TABLE 36: REFLEXES OF * ɔ, *- ɔ̆-, AND *- ɔː-

An	PC	Rade	Jarai	Roglai	Tsat	Chru	Haroi	WCham	PR Cham
---	*-ɔ	-o	-o	-o	-o	-ɔ	-ɔ	-ɔ	-ɔ
---	* ɔ-	-ɔh	-ɔh	-oh	---	-ɔh	-ʋh; -ŭh; -ŭh; -ɔ̆h; -ɔh	-ɔh	-ɔh; -ɔ̆h
---	*ɔ̆-	-ɔ̆ŋ	-ɔŋ; -ɔ̆ŋ	-ok	---	-ɔŋ	-ɔ̆ŋ	-ɔ̆ŋ	-ɔ̆ŋ
---	*ɔː-	-oŋ	-ɔŋ; -oŋ	-oŋ; -oːk	-(u)oŋ	-ɔːŋ	-uːŋ; -uŋ; -ɔŋ	-ɔŋ	-ɔŋ
---	*ɔ̆-	-ɔ̆ʔ	-ɔ̆ʔ	-oʔ	-oʔ	-ɔʔ	-ɔ̆ʔ; -ŭʔ	-ɔ̆ʔ	-ɔ̆ʔ
---	*ɔː-	-ɔ̆ʔ; -ɔ̆k; -ɔ̆ʔ	-ɔ̆ʔ	-oːʔ	-oʔ	-ɔːʔ	-ɔʔ; -ɔ̆ʔ	-ɔʔ	-ɔʔ
---	*ɔː-	-ok	-ɔ̆ʔ	-oːʔ	-oʔ	-ɔːʔ	-ʋʔ; -ɔ̆ʔ	-ɔʔ	-ɔʔ

Not only was PC *ɔ largely borrowed, but in certain environments the vowel was borrowed with a length distinction. The PC *ɔ occurs both long and short before final -ʔ, final-ŋ, and final -k (Tables 37, 38, and 39, respectively).

TABLE 37: *ɔ BEFORE FINAL -ʔ

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ɔ̆ʔ	-ɔ̆ʔ	-ɔ̆ʔ	-oʔ	-ɔʔ	-ɔ̆ʔ	
*ʔakɔ̆ʔ	kɔ̆ʔ	ʔakɔ̆ʔ	akoʔ	akɔʔ	akɔ̆ʔ	head
*gɔ̆ʔ	gɔ̆ʔ; gɔ	gɔ̆ʔ	goʔ	gɔʔ	kɔ̆ʔ	kettle; pot
*chɔ̆ʔ	---	sɔ̆ʔ	choʔ	sɔʔ	---	scoop out
*hɔ̆ʔ	kəhɔ̆ʔ	hɔ̆ʔ	---	---	---	sweat; bleed

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*ŋɔ̃ʔ	---	ŋɔ̃ʔ	---	gah ŋɔʔ (east)	---	above; upgrade
*ʔɔ̃ʔ	---	ʔɔʔ; ʔõʔ	---	ɔːʔ (choke)	---	vomit
*srɔ̃ᵐ	hrɔ̃ʔ	---	---	srɔ̃ᵐ	---	subside
*-ɔːʔ	-ɔ̃ʔ; -õʔ	-ɔ̃ʔ	-oːʔ	-ɔːʔ	-ɔʔ	
*kasɔːʔ	kəsɔ̃ʔ	kəsɔ̃ʔ	kuliʔ soʔ	kəlsɔːʔ	thɔʔ (placenta)	lungs
*kɔːʔ	kõʔ	koʔ -l	koːʔ	kɔːʔ	kɔ̃ʔ; akɔ̃ʔ	white
*mɔːʔ	mõʔ	---	mõːʔ	---	---	wife (mid-)
*ɓɔːʔ	ɓɔ̃ʔ məta	ɓɔ̃ʔ; bõʔ	ɓoːʔ mata	ɓɔːʔ məta	bɔʔ	face cf. nose

TABLE 38: *ɔ BEFORE FINAL -ŋ

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ɔ̃ŋ	-ɔ̃ŋ	-ɔŋ; -ɔ̃ŋ	-ok	-ɔŋ	-ɔ̃ŋ	
*prɔ̃ŋ	prɔ̃ŋ	prɔ̃ŋ	prok	prɔŋ	prɔ̃ŋ	big
*trɔ̃ŋ	trɔ̃ŋ	trɔŋ	trok	trɔŋ	trɔ̃ŋ	eggplant
*salɔ̃ŋ	hlɔ̃ŋ	hlɔŋ	---	---	klɔŋ	forever
*dhɔ̃ŋ	dhɔ̃ŋ	thɔŋ	thok	thɔŋ	thɔ̃ŋ	knife
⧧ *yɔ̃ŋ	yɔ̃ŋ	yɔŋ	yok	---	---	lift; take off
⧧ cadɔ̃ŋ	---	---	caɗok	chəɗɔŋ	cadɔŋ -l	flat basket
⧧ ʔanrɔ̃ŋ	---	---	anrɔʔ -f	---	arɔ̃ʔ	toad
⧧ rɔ̃ŋ	rɔ̃ŋ	rɔŋ	tulaːk turoc	grɔŋ -iʔ	rɔŋ -l	a back
*-ɔːŋ	-ɔŋ	-ɔŋ; -ɔŋ	-ɔŋ; -oːk	-ɔːŋ	-ɔŋ	
*ʔatɔːŋ	tɔŋ	---	atoːk	atɔːŋ	atɔŋ	beat (gong)
*lɔːŋ	lɔŋ	---	---	pərlɔːŋ	---	try, test, prove
*khɔːŋ	khɔŋ (end of rain)	khɔŋ	khoːk	khɔːŋ (dry, sunny)	khɔŋ	dry (weather?)
*rɔːŋ	rɔŋ	rɔŋ	---	---	rɔŋ	nourish
*glɔːŋ -iʔʔ	dlɔŋ	dlɔŋ	dloːk	glɔːŋ	klɔŋ	tall; big; high
*bumɔːŋ -f	mɔŋ (banana stalk)	---	bumoːk -f	---	---	banana blossom
⧧ *kaɗɔːŋ	---	---	kaɗoːk	---	---	get stuck
⧧ *kɔːŋ	kɔŋ	kɔŋ	koːk	kɔːŋ	kɔŋ	bracelet

PC	Rade	Jarai	Roglai	Chru	PR Cham	
⧧ *krɔːŋ	krɔŋ	krɔŋ	kroːk	ia krɔːŋ	krɔŋ	river
⧧ *bɔːŋ	bɔŋ	bɔŋ	boːk	bɔːŋ	pɔŋ	coffin
⧧ hɔːŋ -f	hɔŋ	hɔŋ	hɔŋ -f	hɔːŋ	---	wasp
⧧ jɔːŋ	jɔŋ	jɔŋ	joːk	jɔːŋ	açɔŋ	axe
⧧ lamɔːŋ	emoŋ	rəmoŋ	lumōŋ	rəmɔːŋ	rimɔŋ -i;	tiger
⧧ ramɔːŋ				-r	ramɔŋ -i	
⧧ ʔanrɔːŋ	enoŋ	ʔanɔŋ	anroːk	anɔːŋ	anɔŋ	carry
			-vł			(on a pole)

With the forms in Table 38, the MK influence is particularly clear. Four of the forms have already been analysed as post-Chamic borrowings (⧧), another four are pre-Chamic borrowings (⧧*), and the remainder, while not yet established as borrowings, certainly lack obvious Austronesian etymologies.

TABLE 39: *ɔ BEFORE FINAL -k

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*-ɔ̃k	-ɔ̃k;	-ɔk;	-oʔ	-ɔʔ	-ɔ̃ʔ	
		-ɔ̃k; -ɔ̃ʔ				
*tulɔ̃k	---	təlɔ̃ʔ	---	---	---	disk-like
*pɔ̃k	pɔ̃k	pɔ̃k	poʔ	---	---	to open
*hɔ̃k	hɔ̃k	hɔk	---	hɔʔ	hɔ̃ʔ	pour
	to abort					out; spill
*-ɔːk	-ok	-ɔ̃ʔ	-oːʔ	-ɔːʔ	-ɔʔ	
*dɔːk	dok	dɔ̃ʔ	doːʔ	dɔːʔ	tɔʔ	sit; live; stay
*lɔːk	lok	lok	loːʔ;	lɔːh;	lɔʔ	to peel
		-vf	caloːʔ	lɔːʔ		
*sɔːk	---	---	---	sɔːʔ	---	strike; pound
⧧ kutɔːk	kətuop	---	kutoːk	kətɔːk	---	grasshopper
⧧ kutɔːp	-vf		-f	-f		
⧧ prɔːk	prok	prɔ̃ʔ	---	prɔːʔ	prɔʔ	squirrel

TABLE 40: PC *ɔ APPARENTLY FROM AUSTRONESIAN SOURCES

PC	Rade	Jarai	Roglai	Chru	PR Cham	
*dɔːk	dok	dɔ̃ʔ	doːʔ	doːʔ	tɔʔ	sit; live; stay
*lamɔ;	emo	rəmo	lamo -ⁿ	ləmo	limo;	cow; ox
*ramɔ					lamo	
*bɔh	bɔh	bɔh	bɔh	bɔh	pɔh	fruit;
						egg; classifier

One of the forms containing *-ɔ:k* (*dɔ:k* 'sit; stay; live') is definitely an Austronesian word. The *ɔ* itself also occurs in two more forms apparently inherited from Austronesian sources: *bɔh* 'fruit; egg; classifier for small round objects' and possibly *ramɔl/*lamɔ* 'cow; ox; cattle', although I suspect the latter may be a widespread borrowing instead. However, the remaining forms do not seem to have obvious Austronesian etymologies. Instead, it appears that the overwhelming majority of these forms are borrowings, most of them from MK sources. Certainly, the following are MK forms (see Headley 1976, 1991): 'wasp,' 'axe,' 'bracelet,' 'coffin, casket,' 'grasshopper,' 'river,' and 'squirrel.' Those that were borrowed after the break-up of PC are marked with ⁺. In addition, several more of the forms above, while not identifiable as MK borrowings, nonetheless appear to be borrowings from some source, on the basis of extreme irregularities in patterning (e.g. 'tiger' and possibly 'cow; ox'); the form for 'grasshopper', were it not already identified as a MK borrowing, would still look like a borrowing because of the extreme irregularity of its correspondences. Finally, some other forms look suspiciously non-Austronesian due to their phonetics: the *ɓ-* initial in 'face', the *sr-* cluster in 'subside', the *-nr-* cluster in 'toad', the *ɗ-* in 'flat basket' and 'get stuck'.

4. THE MAIN VOWELS SUMMARISED

It goes without saying that everywhere the details remain to be filled in and clarified. For instance, much can be learned about the first-syllable vowels from a more sophisticated examination of the written records. Similarly, a better understanding of borrowings will contribute to a better understanding of the systemic interactions between the Austronesian and the MK heritage.

Nonetheless, the outlines of the history of PC vowels seem clear. The PC vowel system consists of a core of elements inherited from Austronesian, supplemented and enriched by MK borrowings. In addition, the subsequent reflexes of PC vowels in the various daughter languages is also quite straightforward, with our improvement upon the foundation laid by Lee (1966) and others made possible by an expanded understanding recognition of which forms were borrowings and by a greatly expanded database.

From these patterns we can learn something both about the nature of the earlier cultural contact and about the influence of language contact on vowel systems. The intensity of the early contact between MK speakers and the pre- PC speakers is attested to by the richness of the borrowed component of PC. The effects of language contact are attested to by the restructuring of the original Austronesian disyllables into the iambic morphemes of PC and by the incorporation of a number of new vowel distinctions into the linguistic system.

REFERENCES

Aymonier, Étienne François and Antoine Cabaton, 1906, *Dictionnaire cham-français.* PEFEO 7. Paris: Leroux.

Benedict, Paul K., 1984, Austro-Tai parallel: a tonal Cham colony on Hainan. *Computational Analyses of Asian and African Languages* 22:83–86.

Blood, Doris Walker, 1962, Reflexes of Proto-Malayo-Polynesian in Cham. *Anthropological Linguistics* 4/9:11–20.

Blood, Henry F., 1968, A reconstruction of Proto-Mnong. SIL/University of North Dakota.

Blust, Robert A., 1969, Some new Proto-Austronesian trisyllables. *Oceanic Linguistics* 8/2:85–104.

1980a, Austronesian etymologies. *Oceanic Linguistics* 19/1–2:1–181.

1980b, More on the origins of glottalic consonants. *Lingua* 52:125–156.

1981, The reconstruction of Proto-Malayo-Javanic: an appreciation. *Bijdragen tot de Taal-, Land-, en Volkenkunde* 137/4:456–469.

1983–84, Austronesian etymologies—II. *Oceanic Linguistics* 22-23/1 & 2:29–149.

1986, Austronesian etymologies—III. *Oceanic Linguistics* 25/1-2:1–123.

1989, Austronesian etymologies—IV. *Oceanic Linguistics* 28/2:111–180.

Brennan, Hugh, 1992, Peminjaman leksikal dalam Bahasa Melayu. *Jurnal Dewan Bahasa* 36/6:552–565.

Burnham, E. C., 1976, The place of Haroi in the Chamic languages. MA thesis, University of Texas at Arlington.

Coope, A. E., 1986, *Macmillan's Malay—English English—Malay dictionary.* Singapore: Macmillan.

Donegan, Patricia, 1993, Rhythm and vocalic drift in Munda and Mon-Khmer. *Linguistics of the Tibeto-Burman Area* 16/1:1–43.

Haudricourt, André-G., 1984, Tones of some languages in Hainan. *Minzu Yuwen* 4:17–25.

Headley Robert K., 1976, Some sources of Chamic vocabulary. In Jenner et al., eds 1976, Part 1:453–476.

1991, The phonology of Kompong Thom Cham. In Jeremy H. C. S. Davidson, ed. *Austroasiatic studies: essays in honour of H.L. Shorto,* 105–121. University of London, School of Oriental and African Studies.

Huffman, Franklin E., 1976, The register problem in fifteen Mon-Khmer languages. In Jenner et al., eds 1976, Part 1:575–590.

Jenner, Philip N., Laurence C. Thompson, and Stanley Starosta, eds *Austroasiatic Studies.* Special publication of Oceanic Linguistics No.13. Honolulu, Hawaii: University of Hawaii Press.

Lee, Ernest Wilson, 1966, Proto-Chamic phonologic word and vocabulary. PhD dissertation, Indiana University.

1977, Devoicing, aspiration, and vowel split in Haroi: evidence for register (contrastive tongue-root position). In Thomas et al., eds, 1977:87–104.

Ni Dabai, 1988a, The genealogical affiliation of the language of the Hui people in Sanya Hainan. *Minzu Yuwen* 2:18–25.

1988b, The Kam-Tai of China and the Hui of Hainan. Bulletin of the Central Institute of Minorities 3:54–65.

1990a, The origins of the tones of the Kam-Tai languages. MS.

1990b, The Sanya (=Utsat) language of Hainan island: a living specimen of a linguistic typological shift. MS.

Pittman, Richard S., 1959, Jarai as a member of the Malayo-Polynesian family of languages. *Asian Culture* 1/4:59–67.

Shorto, Harry L., 1975, Achinese and mainland Austronesian. *Bulletin of the School of Oriental and African Studies* 38:81–102.

Smith, Kenneth D., 1972, A phonological reconstruction of Proto-North-Bahnaric (Language data, Asian-Pacific series, No.2). Santa Ana, CA: SIL.

Thomas, David, Ernest W. Lee and Nguyen Dang Liem, eds, 1977, *Papers in South East Asian Linguistics No.4: Chamic Studies. PL,* A–48.

Thomas, Dorothy M., 1963, Proto-Malayo-Polynesian reflexes in Rade, Jarai, and Chru. *Studies in Linguistics* 17:59–75.

Thurgood, Graham, 1996, Language contact and the directionality of internal drift: the development of tones and registers in Chamic. *Language* 71/1:1–31.

1997, Restructured register in Haroi: reconstructing its historical origins. In A.S. Abramson, ed., *Southeast Asian linguistic studies in honour of Vichin Panupong,* 283–295. Bangkok: Chulalongkorn University Press.