On the track of Austric part III. basic vocabulary comparison

La Vaughn H. Hayes

1. Introduction

1.1. Background. In Parts I and II of this series (Hayes 1992, 1997b), it was observed that the primary reason the Austric hypothesis has not been confirmed and Austric generally accepted as an established language phylum composed of the AA and AN language families is because the published supportive lexical evidence has never been considered probatively adequate to the task.[1] This evidence comprises perhaps in toto some 5–600 comparisons, beginning with Wilhelm Schmidt's presentation of 215 sets of putative lexical correspondences between Austroasiatic and Austronesian when he first proposed the Austric hypothesis in 1906. Subsequent contributors (Pou and Jenner 1975, Benedict 1976, Shorto 1976, Diffloth 1977, 1990, 1994, Hayes 1992, 1997a, 1997b) have each adduced lesser numbers of comparisons, with the total of 267 presented by Hayes now surpassing Schmidt's initial contribution. This study will introduce 127 new comparanda and bring the Hayes total up to 394.

Although Schmidt's lexical evidence has apparently never been subjected to a rigorous and comprehensive examination in print, it seems fair to say that its principal weak points are lack of basic core vocabulary comparisons and inclusion of a type of vocabulary easily suspect as being borrowed. Gérard F. Diffloth has essayed a limited review of Schmidt's data, as well as that presented by Saveros Pou, Philip N. Jenner, and Harry L. Shorto, and concluded that "The lexical agreement between Austroasiatic and Austronesian, such as we see it at the moment, is not impressive, whichever lexical class one chooses to look at; but it is undoubtdly there (1994: 312)."

Since the lexical agreement is there, as Diffloth so sagely counsels us, one might think that it remains only to find and present a respectable amount of this evasive lexical material in order to end once and for all the debate about the credibility and validity of the Austric hypothesis. As all investigators, including the writer, have discovered, the matter is not quite so simple as that.

---

1. Abbreviations used here are AA (Austroasiatic), AN (Austronesian), AT (Austro-Tai), CF (composition form). CN (Central Nicobar), E (East, Eastern), FO (Formosan), KY (Khmu' Yuan), MK (Mon-Khmer), MM (Middle Mon), MP (Malayo-Polynesian), MUK (Mường Khến), N (North, Northern), NK (Nyah Kur), OM (Old Mon), P (Proto-), PC (Proto-Chamic), PCEMP (Proto-Central Eastern Malayo-Polynesian), PM (Proto-Mon), PMN (Proto-Mnong), PNB (Proto-North Bahnaric), POC (Proto-Oceanic), PSB (Proto-South Bahnaric), PVM (Proto-Viet-Muong), PW (Proto-Waic), V (vowel), W (West, Western), VN (Vietnamese).

Three reasons for this difficulty can be identified. First, credible AA/AN comparisons are exceptionally hard to find, partly because published AA lexical data have always been in limited supply and partly due to obscuration and concealment of lexical, phonological, and semantic linkages by diachronic change. This fact was pointed out in Austric I and II, and some changes of the obscuring/concealing sort were described in Austric II. Second, many potential AA cognates are so phonologically similar to their proposed AN correspondents that borrowing by Austroasiatic from Austronesian is suspected—curiously, the reverse is never proposed—and whether or not this is actually the case, detractors are able to use the borrowing issue as grounds to discount such comparisons and thereby minimize and/or discredit the available evidence. Third, previous lexical data presentations have arguably not been as comprehensive and systematic as needed to demonstrate convincingly the dimension of lexical correspondence and regularity of phonological, morphological and semantic correspondence necessary to justify a genetic linguistic relationship.

In this part of the Austric series, an attempt will be made to overcome some of the difficulties described above through the means of a presentation of lexical correspondence in the basic vocabulary area. It is hoped that this paper will go far towards conclusively ending the 90-year long debate over the existence of credible lexical evidence for Austric and clear the way to the phylum's general acceptance.

1.2. Purpose and objectives. The purpose of this paper is to demonstrate that a certain degree of lexical correspondence does exist between Austroasiatic and Austronesian in the basic vocabulary area of the lexicon and that this correspondence is of such magnitude and nature that it is more likely indicative of a genetic, rather than a contact relationship. To that end, basic vocabulary reconstructions from Proto-Malayo-Polynesian, one of the two primary AN subgroups—Formosan being the other, will be presented and used as a reference point for discovery of potential lexical correspondents in Austroasiatic.

2. Preliminaries

2.1. The AA and AN phonological systems. The proto-phonemes used in the lexical reconstructions presented in the basic vocabulary comparison in Section 3 are displayed in the following subsections. Since the representations of the PAN and PMP sound systems are identical, Proto-Austronesian is used as a label for both in the display and discussion.

2.1.1. The consonant systems. A major difference between the two systems displayed in Table 1 is the reconstruction of a retroflex series, */T, D, S, Z, N/, for Austronesian.[2] */T, D/ are not accepted by all Austronesianists, and */C/, a

2. In the text, square brackets denote phonetic, slashes phonemic, and curly brackets orthographic representations. In reconstructions, square brackets denote uncertainty, parentheses optionality, and curly brackets proto-forms based on the evidence of a single language.

voiceless retroflex affricate not included in the table, is even more questionable. Other differences are presence in Austroasiatic of */G, x, ɻ, R, [N]/, which have merged with other phonemes or become lost in Austronesian according to Paul K. Benedict (cf. 1975:155). Note that PAA */R/ is a postvelar spirant and not identical to the PAN velar spirant */R/, the AA equivalent of which is represented by */ɣ/.

Table 1. PAA and PAN Consonants

Proto-Austroasiatic						Proto-Austronesian						
p	t	c	k	q	ʔ	p	t	T	c	k	q	ʔ
b	d	j	g	G		b	d	D	j	g		
	s		x		h		s	S				h
	z		ɣ	R			z	Z		R		
	l			ɻ			l					
	r						r					
w		y				w			y			
m	n	ñ	ŋ	[N]		m	n	N	ñ	ŋ		

2.1.2. The vowel systems. The PAN vowel system comprises four phonemes: */i, ə, a, u/ (*/ə/ is customarily represented by */e/). The PAA vowel system is provisionally reconstructed with six sounds: */i, e, ə, a, u, o/, but it is already becoming clear that additional vowel phonemes and diphthongs will probably have to be reconstructed eventually. Until this revision is possible, PAA vowel reconstructions must be regarded as very tentative work in progress.

2.2. Basic vocabulary

2.2.1 Basic vocabulary defined. The concept of a basic vocabulary comprising the most commonly used words of a language is probably as old as the study of language itself. It is apparently only during the past five decades that the concept has taken on a specialized time-linked meaning. This development can be traced back to Morris Swadesh who introduced in 1949 a lexicostatistical method he called glottochronology (cf. Lehman 1992:175), which was based on the assumption that the most common words of a language—its *basic core vocabulary*—are retained over time at a precise rate. In this context, Jacques B.M. Guy (1995:63) defines basic vocabulary as follows:

> "Basic vocabulary", as understood in glottochronological theory, is then merely "stable vocabulary", ideally the set of universally elicitable lexical items such that each member of the set has the same high probability of surviving unreplaced through a unit period of time as any other member. Even if such a set were not likely to be empty, and if there were a way of computing the retention rates of individual items without carrying out a preliminary classification of the languages involved, its membership could only be determined after examining

> the whole vocabularies of all languages. Relaxing the requirements for inclusion into the set (using items elicitable in most communalects of a given language group, with approximately the same individual retention rates) would still require the preliminary collection of very large vocabularies from many languages and dialects, making glottochronology based on lexicostatistical data an impractical proposition.

In recent years, glottochronology has been generally discredited because Guy and other scholars have demonstrated that no such precise rate of diachronic retention can be shown to exist for all natural languages or even for all members of a given natural language family. The concept of a basic vocabulary which is resistant to diachronic replacement and thus more stable over time than the rest of the lexicon remains nevertheless useful to historical linguists. This is so because such vocabulary would tend to retain evidence of the lexical, phonological, and morphological characteristics of the ancestral language longer and more faithfully than other lexical categories and thus provide a greater degree of evidential and probatory proof of the ancestral language's nature and any genetic relationship existing between the languages being compared.

2.2.2. Basic vocabulary determination

2.2.2.1. General. As Guy points out, determining the content of a universal basic vocabulary would require ideally classification of the vocabulary of all the world's languages or as a minimum classification of the vocabulary of all the languages in the specific language family under study. Such classification is not yet possible on a world-wide scale, nor is it feasible in many language families, to include Austroasiatic and Austronesian, because the lexicons of all member languages have not yet been recorded in sufficient detail or at all. Those facts have not hindered linguists from compiling basic vocabulary lists.

Swadesh constructed a number of basic vocabulary lists containing varying numbers of common words; his 200-word list is given in Lehman 1992:180f. Other scholars have also prepared such lists. Some very short ones (6–20 words) may have universal applicability, but it appears that as a rule of thumb, the more words placed on a list, the less appropriate it becomes for an increasing number of languages. Some words must be invariably added or subtracted in order to reflect more accurately the common word situation in specific language groupings.

2.2.2.2. AA basic vocabulary. Linguists of the Summer Institute of Linguistics have designed two lists for usage with the MK languages, both modifications of the Swadesh 200-word list, one comprising 207 words (Thomas and Headley 1970: 411–16), the other 281 (Miller 1994:72–81). The modifications are based on lexical compilations from a subset of the MK languages; they are not the result of a general lexical classification of Mon-Khmer, an as yet impossible project due to lack of the necessary lexical materials on all MK languages. To the writer's knowledge, no basic vocabulary list has ever been prepared for the Munda languages.

As a result of this situation, it is not possible to present in this study a word list with any bonafide claim to representing the AA basic vocabulary, either contemporarily or at any earlier historical stage of this language family. Constructing such a list, an ad hoc effort at best, would doubtlessly require more time and effort than expended on this entire presentation.

2.2.2.3. AN basic vocabulary. In general, AN lexical and historical studies are far advanced in comparison to those in the AA field. Although an all-language-inclusive lexical classification is no more possible than it is in Austroasiatic, AN lexical compilations of all sorts, from word lists to dictionaries, are generally more numerous and more comprehensive than in Austroasiatic. On the historical side, reconstruction of high-level proto-languages, to include Proto-Austronesian, is fait accompli in the AN field whereas in Austroasiatic, no comparable reconstructions exist. Accordingly, Austronesianists are able to compile lists of basic vocabulary reconstructions which can claim (with a degree of accuracy unknown to the writer) to reflect the common words of earlier language stages in the AN hierarchy.

No such list exists for the PAN level, but Robert Blust (1993:280–4) has prepared a list of PMP basic vocabulary reconstructions, which will serve the purposes of this study. This is a 201-word AN-specific modification of the Swadesh 200-word list; however, Blust cites more than one proto-form for certain English glosses, affixed derivatives for two, and a homophone for two others, which raises the list's word total to 240. In this presentation, the derivatives have been removed and the homophones combined, reducing the total to 237 (see Section 3).[3]

In the referenced source, reconstructions are also simultaneously listed for the primary descendants of Proto-Malayo-Polynesian, i.e. the CEMP, Central MP, and Oceanic proto-languages. These lists are not reproduced here; however, where possible AA correspondents exist for proto-forms found on these lists, but not on the PMP list, these comparisons are cited in Section 3 as additional items.

2.2.3. Basic vocabulary comparison. Optimally, reconstructed basic vocabularies for the PAA and PAN levels should be compared in order to determine what lexical, phonological, and morphological correspondence exists between these proto-languages and how this correspondence affects the question of their genetic relatedness. As noted in Subsections 2.2.2.2 and 3, however, no PAA or PAN basic vocabulary exists, and it is not possible to assemble one or both within any reasonable amount of time due to lack of supporting lexical materials. Hence, a different and less optimal approach must be taken here.

Accordingly, the PMP basic vocabulary reconstructions presented in Section 3 are used as a referential base. Potential lexical correspondents to those proto-forms are adduced from Austroasiatic, if such correspondents can be found, together with their provisionally reconstructed PAA antecedents. The rationale underlying

---

3. The excised derivatives are PMP */b-in-ahi, ba-b-in-ahi/ 'female/woman' and */h-in-ipi, h-um-ipi/ 'dream'. The combined homophones are PMP */qasawa/ 'husband' and */qasawa/ 'wife'.

this methodological approach is simple and not believed to be theoretically inconsistent.

Since basic core vocabulary is resistant, but not invulnerable to diachronic replacement, it follows that some basic vocabulary items will be retained and some will disappear, but disappearance from the basic vocabulary category need not mean that an item has been lost from the lexicon; it may have simply undergone semantic shift and still be there. As a consequence, some subset of the non-basic vocabulary in some languages will correspond to a subset of the basic vocabulary in other languages in a set of genetically related languages. More importantly, the former will retain the same characteristics found in the latter which are useful to historical comparatists as noted in Subsection 2.2.1, and recovery of such lexical correspondences should provide comparisons of historical value equal or comparable to comparisons found only in the basic vocabulary.

2.3. Methodology

2.3.1. Some basic issues. Historico-comparative studies of languages such as the Austroasiatic and Austronesian, the genetic relatedness of which has not been verified, are generally known as megalo- or long-range comparisons, and critics of such comparisons routinely point to an acclaimed lack of rigor in the methodology applied in such endeavors. Although sometimes overdone, the criticism is doubtlessly appropriate in one respect. In conventional methodology, the bottom-up approach is used, whereby comparison is begun at the lowest level of the taxonomic hierarchy and gradually stepped up to the top-level. At each level, proto-languages are reconstructed and used as the basis for the next-step comparison. Due to the general lack of historical studies and reconstructed proto-languages mentioned in Subsection 1.1, a top-down approach must be used in the Austric comparison, whereby data from the lowest level AA languages and such mid-level reconstructions as are available are compared and the AA proto-language reconstructed on that basis. The PAA and PAN reconstructions are then compared in order to gain insight to the nature of Proto-Austric.

Although the top-down approach does not seem to be theoretically unsound, it should probably be judged as methodologically flawed because some quantity of the diachronic developments which occurred between the top and bottom levels probably cannot be recovered and this inability introduces an indeterminate degree of inaccuracy and unreliability into the results of the comparison. Whether or not the comparatist is willing to accept this risk, is of course a subjective decision. In the writer's case, it is felt that the long-term advantages accruable from taking this approach to AA and Austric historical studies will outweigh the disadvantages of some limited inaccuracy and purist disapproval.

2.3.2. Proof of genetic correspondence. The general purpose of the writer's series of papers on Austric is presentation of linguistic evidence supportive of Schmidt's Austric hypothesis, and its objective is affirming that Austric is a valid linguistic construct. As stated in Austric I (Hayes 1992:148), multiple agreement in the basic core vocabulary is one type of evidence considered proof of the genetic relationship

which must exist between Austroasiatic and Austronesian if the Austric hypothesis is to be verified. But a simple listing of lexical comparisons seemingly evidencing such agreement is not sufficient to accomplishing our objective; these comparisons must also exhibit genetic 1) phonological and 2) morphological correspondence, defined in Austric I as 1) recurrent correspondence of phonemes as to their position in phonemic strings expressing morphemes with correspondent semantic contents and 2) phonemic correspondence of any part of phonemic strings expressing word forms, but only of such parts having the same position with respect to the other parts of the words compared.

The criterion by which demonstration of such correspondence is considered rigorous is apparently that no exceptions are allowed. That is, all correspondences in the lexical comparanda presented must be recurrent or regular and the semantic contents of the comparanda must be identical. Unfortunately, that criterion cannot be met across the board at this stage of the AA/AN comparison. In truth, it is rarely, if ever met, in any historical comparison; some unexplanable irregularities almost always exist. In conventional comparisons, that fact can be conveniently swept under the rug because the irregularities are few, the regularities in majority; in long-range comparisons, the irregularities are often numerous and unavoidable and thus must be openly displayed and faced by the comparatist.

In the PAA reconstructions presented in Section 3, most phonemic correspondences are in fact regular, at least where the consonants are concerned. The vowels are a different matter, as indicated in Subsection 2.1.2. Many phonemic correspondences which would otherwise appear to be irregular are explanable in terms of environmental phonetic or morphophonemic conditioning. For example, AA */b/ correlates regularly to AN */b/, but the /m/:*/b/ correspondence in the Brou {ramurl}/PMP */Rabun/ 'cloud' comparison in Subsection 3.2 appears to be irregular. However, it becomes clear in the analysis that */b/ must have alternated with */mb/ in the distant past and the cluster coalesced as */m/ in some cases, as in Brou {ramurl}, but remained */mb/ in others, as in Mundari {rimbil} 'cloud', or was simplified to */b/, as in PMP */Rabun/, cf. AN */Ra(m)bun/. This alternation apparently resulted from usage of an ancient affix */N/ which could be optionally inserted in the word (cf. Hayes 1992:167ff.). Dialectal variation probably explains why this affix accreted to */b/ and the cluster coalesced in selective fashion. Such explanations are not unknown in methodologically conventional presentations.

Even in conventional comparisons, the requirement of identical semantic content is not always rigorously applied, but the degree of laxity permitted is usually not large. The same approach is used here. At least one cognate form expressing identical content is used where available in the lexical comparisons. When none is available, the comparanda listed must express a meaning which is minimally divergent from or logically translatable into the meaning of the PMP form. In Subsection 3.8, for example, the meaning of VN {ước} 'to desire, wish for, hope for' is correlated to that of PMP */ma-pia/ 'good' under the assumption that what one desires or wishes for is usually something good. This assumption is less risky than it might seem because the corresponding AN proto-form */pi[y]a[h]/ means 'desire, desiderative marker', PMP */ma-pia/ being a stative derivative.

In recognition of the fact that the justification for a number of the correlations seen in Section 3 may not be clear, liberal usage of explanatory footnotes is made.

2.3.3. Identification of loanwords. In Subsection 1.1, the problem of AN loanwords in AA comparative data was briefly addressed. It is a methodological necessity that all such words be identified and eliminated from the basic vocabulary comparison in order that an accurate evaluation of the AA/AN lexical correspondence can be made. This task is relatively easy when the AA and AN forms are phonologically and semantically very divergent or exactly identical, as in the case of Semelai and Malay {bulan} 'moon', cf. PMP */bulan/ 'moon' in Subsection 3.2. It becomes more difficult when the AA form is minimally divergent, phonologically and/or semantically, from its PMP correspondent, as in the case of Katu {tam}, PMP */ma-qitem/ 'black' in Subsection 3.8. Here, criteria for loanword identification must be developed, and thus far, three primary ones have been established.

1) Presence of a Munda cognate is grounds for inclusion of a comparison. The rationale is that there is no known reason to suspect that Munda was ever in contact with Austronesian; hence, it cannot possess any AN loanwords. In the case of the 'black' comparison, Sora {süm} 'grope in the dark' appears to be a reflex of AA */itəm/ > */cəm/. Katu {tam} may still be an AN loanword, although there is little reason to think so in view of the fact that Chamic possesses no reflexes of PMP */ma-qitem/, but even if it is, that is not grounds for rejection of the entire comparison due to the Munda cognate's presence.

2) Presence of a distinctive phonological correspondence is also grounds for inclusion. A correspondence of this type is /l/:*/n/, as in the comparison Bahnar {lai} 'mound of dirt', PMP */qenay/ 'sand' in Subsection 3.2. The Bahnar form cannot be borrowed from AN because Austric */-l-/ > AA */l/, AN */n/ and a later change of this */n/ back to */l/ is unknown in any AN language with which Bahnar could ever have been in contact.

3) Presence of a distinctive phonological development is further grounds for inclusion. This is a development which makes it clear that the diachronic evolution of the AA lexical form has been so different from that of its AN correspondent that the AA form could not possibly have been borrowed from Austronesian, as in the case of Khmu' {?om}, PMP */danum/ 'water' in Subsection 3.2. Khmu' /?om/ continues the Austric root */[?]om/, while PMP */danum/ reflects an affixed derivative, */zal[?]om/, of the same root. Katu {dalóm} 'gulp' is apparently a reflex of the same Austric derivative, though it may have been created at a later date. Comparisons in which Austroasiatic has retained a reflex of the bare Austric root and Austronesian a reflex indicating affixation of the same root or its compounding with another root are fairly common. Much less often, it is Austronesian which has retained the bare root and Austroasiatic the morphologically more complex lexical form.

3. Basic vocabulary comparison

3.1. General. The following 12 subsections are titled according to semantic category: nature, flora, fauna, anatomy, kinship, cultural artifacts, descriptives, verbs, pronouns, numerals, prepositions, and miscellaneous. The numbers given after the subsection title, such as (29/22), denote before the slash the number of PMP forms in the category and after it the number of PMP/AA comparisons presented.

In the PAA column, a single asterisk preceding a proto-form identifies it as valid for Proto-Austroasiatic, two asterisks identify it as valid for Proto-Munda or Proto-Mon-Khmer only. In the AN column, PC reconstructions are cited after the PMP forms when available. Most of these come from Lee 1966:179–220, with the remainder provisionally reconstructed by the writer on the basis of Lee's findings.

3.2. Nature (29/22)

Austroasiatic	PAA	Austronesian
Pacoh abóh, Chrau vuh 'ashes', Bonda bu? 'to smoke'	*qabuh	PMP *qabu, PC *habəu 'ashes'
Mundari (V284)[4] rimbil, Brou ramurl, MUK mâl 'cloud'	*γ[a]mb[o]l	PMP *Rabun 'cloud'[5]
VN (*[k]ra:w >) sao 'star', Katu charo 'polar star'	**[j](a)raw[6]	PMP *qalejaw 'day'
Bonda tubɔk'/tubuk' 'earth', Chrau võq 'mud', Mon khabuik 'fine powder or dust'	*buk	PMP *qabuk/qapuk 'dust', PC *buk 'mud'
Mundari (V403) ɔte 'soil, earth, field, land', PW *ktɛ?, PM *tĭ? 'earth'	*teq[7]	PMP *tanaq/taneq 'earth/soil', PC *tanãh 'earth'
Sora (*[s]oy >) oy 'kindle',	*[s]uy,[8]	PMP *hapuy, PC *?apui 'fire'

4. Lexical data cited in Pinnow 1959 are referenced by his indicial numbers, V denoting those found in the section on vowels, K those in the section on consonants.

5. Cf. AN */Ra(m)bun/ 'atmospheric obscurity/turbulence, atmospheric cloudiness', AT */γ[a](m)bon/ 'rain, drizzle (fine rain), mist, snow, clouds, dusk, dark, sky', Proto-Kam-Sui */ɓun[1]/ 'sky', Proto-Mjuenic */?buən[C]/ 'cloud'.

6. Limited to Vietic and Katuic. The original meaning of the root was perhaps 'sun', whence 'day' via 'light of the sun' or the like. Such Bahnaric forms as {srŏ} [srɔ?] in Jeh {srŏ ay} 'dawn' may also belong to this etymology. Also cf. AN */ha(n)daw, qa(ŋ)jaw/ 'day, sun', AT */(q)a(n)draw/ 'sun, star, sky', Proto-Kam-Sui */?dra:u[1]/, Proto-Thai */?d[l,r]au[A]/ 'star'.

7. This comparison is based on the assumption that the PMP form reflects an infixed derivative of the Austric root, */taneq/ or possibly */taleq/ in view of Bunun /dalaq/ 'earth'.

8. The root may have been */suy/ (or */uy/, cf. Koho {ôs} 'fire'), with */sa(m)puy/ an infixed derivative. Note that */s/ > /0/ in Sora and */s/ > */h/ > /0/ in PMP, cf. Atayalic */sapuy/ 'fire' in Formosan. The Chamic glottal stop is secondary, appearing before all vowels. Also cf. Kharia {sului} 'get hot', which may confirm the AA initial's identity as */s/.

Pearic puy 'tinder', Katu mpoih 'fire'	*[sa](m)puy(s)	
Sora (V384) umod-ən 'fog, mist', Khmu' (hm)puut 'clouds, fog'	*(m)put[9]	PMP *kabut 'fog/mist'
Khmu' rɨyh 'forest without underbrush', Pacoh arurih rứng trứng 'wilderness', Sengoi jeres 'the jungle'	**[a]l̥ac(i)[10]	PMP *halas 'forest'
Chrau tanlô 'lake, pond'	{**ləw}[11]	PMP *danaw, PC *danãu 'lake'
Santali bilit, Nicobar pələ-leɛt, Pearic la:t, la:c 'lightning'	*(bi)l̥at(i)	PMP *kilat 'lightning'
Katu kamlaaq, Khasi leilieh [leɪlie?] 'lightning'	**(bi)[l̥]aq[12]	PMP *qusilaq 'lightning'
Katu baraal 'pale', Bateg Deg bəyl̥ (*r > y) 'white', Bahnar mơnhal 'very bright light or sunshine'	**b(i,a)l̥al[13]	PMP *bulan 'moon', PC *bilān 'month, moon'
None[14]		PMP *beRŋi/beRŋin 'night'
Brou cuyal, Old Mon kyāl, Khmer khya'l 'wind'	**ka[ñj]al[15]	PMP *quzan, PC *hujān 'rain'

---

9. Limited to Sora and Khmu', and the Sora development is obscure. Suspected Aslian reflexes, such as Sengoi {kabus} 'mist, fog', mirror Malay {kabus} 'dark, misty', cf. also Malay {kabut} 'dark', and are probably loanwords. Cf. also AT */[ka](m)bot/ 'mist, cloudy' and Proto-Thai */ʔbot/ 'cloudy'.

10. Also cf. Mundari {ɑrɑ:} 'plants, vegetables', Pareng {ərɑ:} 'a tree, wood', and Sora {ərɑ:jɑ:ŋ-ən} 'forest', which may reflect an underlying Austric root */[a]l̥a/, whence */[a]l̥at(i), [a]l̥ats(i)/, or a truncation of the cited AA proto-form.

11. The Chrau word appears to be an isolate in Austroasiatic, but in view of the lateral, it is difficult to assert that it is borrowed from Austronesian. VN {ao} 'pond' is possibly reflective of an Austric root */[ʔ][a,ə]w/, but suspect as a loan, cf. Ancient Chinese */ʔau/ 'bay, bank', Thai {ʔāw} 'gulf, river bend'.

12. Limited to Katu and Khasi, and the Khasi phonology is obscure, but Chamic has apparently borrowed this word from Mon-Khmer, cf. NRoglai {chacalàq liaq} 'lightning' with a glottal stop (orthographic q) instead of /h/, the regular Chamic reflex of PMP */q/. PM */ktah/ 'lightning strike, struck by lightning' may reflect a variant, */plaq/, in which the cluster coalesced as */t/.

13. This comparison is based on the assumed semantic shift of '(pale) white' to 'moon light' to 'moon'. The Austric root may be reflected in Jehai (Sem. Jarum, C184) {āl} 'cloud' and Kintaq Bong (Sem. Buk. Max., C184) {al} 'sky'. Khasi {blar-blar} and {blir-blir} 'pale' appear to reflect metathesis of variants */baral/ and */biral/. Also cf. AT */(q)(m)bul̥al/ 'white, moon, silver', Proto-Thai */ʔb[l,r]ïen/ 'month, moon'. Malay {bulan} 'moon' and reflexes of PC */bilān/ have been borrowed by a number of MK languages in Vietnam and Malaysia.

14. AA forms such as PNB */măŋ/ 'night' suggest */m[b]əŋ,/ a possible correspondent of AN */beŋ[ih]/ 'night'; however, this comparison cannot hold up to scrutiny in face of the unevidenced */b/ in Austroasiatic and the post-Dempwolff revisions of the AN proto-form to */beRŋi[ʔ,h]/.

15. Cf. AT */(q)[u](n)dzal/, EFO */qudzal/ 'rain'. The Austric root may have been */zal/, reflexes of which some AA dialects apparently retain, cf. Pearic {khsál}, Chrau {chhal} 'wind'. Other dialects reflect the developmental sequence */ndz/ > */n'j/ > */ʔy/ > /y/, whereby the ancient cluster became a glottalized or

Pareng bɑsut, Kuy pos, Jehai mpɔj 'salt'	*qa(m)puc(i)[16]	PMP *timus 'salt'
Katu riêq 'food has salt', Sengoi gerak 'salty (to taste)', Mal suak 'salt'	**si[ɣ]aq, **su[ɣ]ak[17]	PMP *qasiRa, PC *sara 'salt'
VN lây 'miry, swampy, marshy', Bahnar lai 'mound of dirt', Nyah Kur lɛɛ, rèɛ 'ore, mineral'	**lay	PMP *qenay 'sand'
None[18]		PMP *tasik 'sea/saltwater', PC *tasi? 'sea'
None[19]		PMP *laŋit 'sky', PC *laŋĩ? 'heaven, sky'
Sora poro: 'become smoky', mor 'spread as smoke', Stieng pôr 'smoke out of a hollow tree'	*(m)pə̤[20]	PMP *qebel 'smoke'
Pearic (mə)čok 'to smoke', Chrau nhŭq 'smoke', Bahnar yuh 'steam, smoke'	**(ñ)cuk, **(ñ)cuq[21]	PMP *qasu 'smoke'
Bonda mərtək', Sora tuj-tuj-ən 'star', PM *kntuu? 'moon, month'	*t[o]q(i)[22]	PMP *bituqen, PC *pitŭ? 'star'

imploded cluster which dissimilated and reduced to the palatal glide.

16. The South Munda glosses indicate */b[u]cut/, perhaps an old compound or suffixed deriva- tive, cf. also Sora {bəsud-ən} 'salt'. Other suspected AA members of this etymology indicate */(m)puq, puluq/, etc., possibly suggesting an Austric root */pu/, whence */(m)puts/ > AA */qa(m)puc(i)/ and AN/PMP */timus/.

17. Katu apparently reflects development of one variant, */si[ɣ]aq/ > */səriaq/ > {riêq} [riə?], which could be an AN loan, but not from Chamic, while Mal reflects another variant, */su[ɣ]ak/ > */sruak/ > */śuak/ > {suak}, where the velar is another suffix or perhaps a reflex of */c/ from earlier palatalized */q/, cf. also */su[ɣ]aq/ > */sərua?/ > High Katu {haruôq} [həruə?] 'tasteless', i.e. 'unsalted'.

18. Bahnaric and Katuic have borrowed this word from Chamic, cf. Katu {tachiiq}, Jeh {dak sĩ} 'ocean'.

19. AA */(p,b)leŋ(i)/ 'sky' was cited in Hayes 1997b as corresponding to AT */[ndu]ḷaŋ[it]/, whence AN (and PMP) */laŋit/ 'sky' per Benedict; however, the AA form may correspond more correctly to AT */b[a]luŋ/ 'sky, rain, thunder', which has reflexes in Formosan, cf. WRukai */ta-buluŋa(n)/ 'sky', but not PMP. Sora (V364) {rɑŋge} (CF {riŋ}) 'air' possibly corresponds to AN/ PMP */laŋit/, but it is an isolate and the phonology is obscure. The AA term for 'sky' may have been */[l]u[ʔ]aŋ/, cf. Khmu' {ləwaaŋ}, Chrau {lawang}, and Sora {ru·ɑ:ŋ} 'sky', which correspond to AN */hawaŋ/ 'air, atmosphere, sky'. This reconstruction suggests an Austric root */[ʔ]aŋ/ which may be incorporated in both AN */laŋit/ and */hawaŋ/.

20. Limited to South Munda and Stieng.

21. The */ñc, ñj/ clusters evolved to /j, ñ, (?)y/ under imprecisely known circumstances, cf. also Sengoi {nyu} 'give off smoke', Khmer {ja'k} 'suck in, smoke (tobacco)', Alak {yɔɔk hiit} 'smoke tobacco'. Also note Proto-Philippines */qasuk/ 'smoke'.

22. The MK reflexes could be borrowed from Chamic on phonological grounds, but most of them mean 'sunlight' or 'sunshine', suggesting that they are not recent loans from Chamic.

Khasi máw, PW *smo?, PM *tmɔɔ? 'stone'	**tamuq[23]	PMP *batu, PC *patəu 'stone'
Kharia g'aṛay, Khmer phgar, PM *gər-gər 'thunder'	*gər[24]	PMP *gurgur 'thunder'
None		PMP *kudug 'thunder'
None		PMP *wahiR 'water (fresh)', PC *?ia 'water'
Khasi um, Khmu' ?om 'water', Mundari (K542) um 'wash the body, bathe'	*[?]om[25]	PMP *danum 'water (fresh)'
None		PMP *haŋin, PC *?aŋĩn 'wind'
None		PMP *taqun, PC thŭn 'year'

3.3. Flora (9/8)

Katu (High) asɑq 'leaf', Sengoi caak [ca:?] 'branch'	**zaqa[26]	PMP *daqan 'branch', PC *dhān 'bough'
Sora puŋ-puŋ 'form spikes, ears or heads', VN bông 'flower', Pacoh apong 'ear of corn, tassel of banana'	*(m)puŋ[27]	PMP *buŋa, PC *buŋã 'flower'
Kensiu kəbə?, Sabum kəmɔ? 'fruit', Stieng moq 'type of small fruit'	**(m)b[o?a]q	PMP *buaq 'fruit', PC *boh 'egg, fruit, ball'
Sora ji:ŋ-ən 'weeds', Chrau canji 'grass, weeds', Stieng cɑnji 'overgrown with weeds'	*(n)je[28]	PMP *baliji 'grass'
None		PMP *udu 'grass'

---

23. This comparison is based on the assumption that either metathesis, as in Austric */tabuq/ > AN */batuq/, or infixation, as in */batuq/ > */batamuq/ > AA */tamuq/, has occurred. Austric or AA */tambuq/ > AA */tamuq/ is also possible.

24. Well represented in Austroasiatic, but more often in forms meaning 'drum' than 'thunder', cf. also AT */(ŋ)guruq, gərəq/ 'rumble, growl, thunder, snore'.

25. Reflexes are found pan-Austroasiatic, but the meaning 'water' is retained only in Khasi and Khmu'; elsewhere, semantic shift to a variety of things and activities related to 'water' has occurred. Chamic has no reflexes of PMP */danum/, which probably reflects Austric */zal[?]om/, but cf. PC */minum/ and AN */hinum/ 'drink' which are based on the same root.

26. Poorly evidenced and somewhat speculative. The Austric root may have been */qa/, whence */z(i,a)(n)qa(n)/, cf. Mon {saka} /həka?/ 'toothstick, twig or slip of wood for cleaning teeth', High Katu {ajuruh} 'snag', Korwa (V274) {siŋ}, Khasi {diŋ} 'tree', Bahnar {dɔng} 'branch', the diverse phonological developments apparently reflecting divergent stress shift patterns.

27. Katu {bangura} 'flower, tree' is the only AA form evidencing the second syllable seen in AN */buŋa[?,h]/, and it is probably borrowed from Chamic. If the comparison is valid, then the second syllable must be assumed to be either a suffixal construction or the reflex of another root in compound with */puŋ/.

28. Limited to Sora and South Bahnaric.

Danaw kătən[4], kătŏn[4], Kotua tuan, Bahnar don 'ear'	**da[q]un[29]	PMP *dahun 'leaf'
Mundari (V366) uṛ 'bark of a tree', Palaung hu:r(-ɛ), Nha Heun sruat 'skin'	*(s)uɣ(at)[30]	PMP *uRat 'root', PC *ʔurăt 'vein, tendon, nerve'
Mundari (V366) harta 'bark of the tree', Sora ə-kɑr 'nail(s) (of finger or toe)', Jah Hut jaŋkar 'root'	*(n)qaɣ[31]	PMP *wakaR 'root'
PW *kho? 'tree', PM *chuu? 'tree, wood', Semai jəhu? 'tree'	**k(i,a)hi(uq)[32]	PMP *kahiw 'wood', PC *kaiəu 'tree'

3.4. Fauna (16/10)

Black Riang luk[2] ak[1] 'crow', Katu plok 'pigeon', Mon gacem k(a)lok 'owl'	**luk	PMP *manuk 'bird', PC *manũ̃? 'chicken'
Bonda gusɔ?, PW *sɔ?, VN chó 'dog'	*cu(q)	PMP *asu, PC *ʔasəu 'dog'
None		PMP *qali-wati 'earthworm'
None		PMP *kalati 'earthworm'
Juang susuter(ɔ), Pacoh tireal tiröl, PVM *t(ə)lur? 'egg'	*teḷoR[33]	PMP *qateluR/qiteluR 'egg'
None		PMP *bulu, PC *baləu 'feather'
Santali (V64) hako, PW *ka?, PM *kaa? 'fish'	*(n)qa(q)[34]	PMP *hikan, PC *ʔikan 'fish'

29. Limited to Palaungic and Bahnaric. Note that */d/ > /t/ in Danaw and Kotua. The MK forms are suspected remnants of an old compound meaning 'leaf [of the] ear', cf. Malay {daun tĕlinga} ('leaf' + 'ear' =) 'outer ear'. In view of their locations (Bahnaric in the central highlands of southern Vietnam, Palaungic in northern Burma), no reason exists to suspect borrowing from Malay.

30. This comparison is based on the assumption that PMP */uRat/ is a reflex of AN */huRat/ 'sinew, tendon, vein'. The initial sibilant was apparently a prefix, for */s-/ is normally retained as /s-/ in Mundari (*/s-/ > Palaung /h-/). The original meaning of the root is unclear, but may have been 'skin', whence 'bark (= skin of tree)', 'sinew (= strip of skin)', and the like.

31. Jah Hut {jaŋkar} 'root' could be borrowed from Malayan or Aceh, but this seems unlikely in view of Diffloth's PM */cŋkɔɔr/ '(Nyah Kur) exterior envelope of vegetals, outer skin of coconut, tree bark, (Mon) bark of tree, fibrous outer husk of coconuts'.

32. The phonological development is obscure, but the Austric root was perhaps */hi/ or */he/ suffixed with */u/ and/or */q/, cf. also Palaung (Panku) {her[2]} 'tree, wood' and PMN */čhi/ 'tree'.

33. Replaces */[q]]iɣu[]]/ cited in Hayes 1992:159. No reflexes found in Chamic, but cf. Malay {tĕlor} 'egg'. Limited to Katuic and Vietic in Mon-Khmer; the Juang form is an isolate in Munda, hence a speculative member of this comparison. Bonda {lɔr} 'plantain to appear on the tree' and Sengoi {lor} 'to bud, sprout (of a plant)' may be reflexes of the underlying root.

34. Certain Munda reflexes indicate */qa/, whence Santali {hako} ({ko} is a plural or class suffix). Since the voiceless postvelar generally shifts to */k/ after a nasal, a prefixed form, */nqa/, must be recognized as the source of the MK reflexes (and presumably the AN as well).

Katu kóót 'lice, fleas, bugs', Khmer saṅkœc 'bedbug, flea', Chrau sicăch 'tick'	**k[o]t(i)	PMP *kutu 'head louse', PC *kutəu 'louse'
Bahnar kɔmŏk 'a type of very small bee-like insect', Thavung məyh, Khmer mūs 'mosquito'	**muk(i)	PMP *ñamuk, *lamuk, PC jamŭ? 'mosquito'
Bonda gubu 'a kind of rat', Riang (Black) kă̆bu[1] 'rat, mouse', Mah Meri (Bes. K. L., R33)[36] kanē' rĕbū' 'mouse'	*b[əw][35]	PMP *labaw 'rat'
Kharia lur, Sora lo?or (CF lor) 'a kind of snake', Bahnar 'bih tep-lar 'a very small snake that is extremely poisonous'	*[su][ļ]aR[37]	PMP *ulaR, PC *?ula 'snake'
None[38]		PMP *nipay 'snake'
Pacoh tuvai apiang 'spider web', Bahnar wai 'spider', Pearic təkliw 'butterfly'	**wa(i,q), **k[i]law[a][39]	PMP *lawaq 'spider'
None		PMP *ikuR, PC *?iku 'tail'
Sora kɑppɑ:-n, Mah Meri kəmphɛk 'wing', Khmer pa'k 'shake, stir, agitate, wave, wag'	*ka(m)pak	PMP *kapak 'wing'
None		PMP *panij 'wing'

3.5. Anatomy (23/20)

Boriwen kúat 'back', Halang kuyq 'small of the back of the	**ko[d(i)][40]	PMP *likuj 'back'

35. Mah Meri {rĕbū'} looks like an AN loanword, but no reflexes of PMP */labaw/ have been found in Chamic and Malay.

36. Parenthetical identifications after the names of Aslian languages consist of the names and indicial numbers used with data found in Skeat and Blagden 1906. The language name correlations are taken from the concordance presented in Benjamin 1976:125–6.

37. Replaces */u[ļ]əγ(i)/ cited in Hayes 1997b:28. The meaning 'snake' is poorly represented in Austroasiatic, and it appears likely that 'snake' was not the original sense of the Austric root. That meaning may be reflected in OM {-lor} 'be supine (?) [hypothetically slor]', from which Spoken Mon {lo} 'lie supine, (snake) wriggle', and (*/glar/ >) {glaw} 'lie on one's back'.

38. Also cf. Bahnar {'bih} and PMN */bĭs/ 'snake', which may be reflexes of */mpays/ or the like, but cannot be presently disambiguated from reflexes of PAA */(um)pac(i)/ 'poisonous snake', cf. PMP *upas 'poisonous, as a snake'.

39. The root was presumably */wa/, whence inflected forms */wai, lawaq/, etc. It seems unlikely that Mah Meri (Bes. K.L., B479) {awā'} 'large butterfly' and Mendriq (Pang. Gal., S378) {tawâh}'spider' were borrowed from Malay {laba laba} 'spider' or Acheh {gĕlabah} 'cobweb'.

40. Semantic shift from 'back' to 'upper back' to 'back of neck or head' to 'head' is evidenced in the MK languages which possess reflexes of this word. The glottalized {yq} [y?] final occurs only in the two closely related Bahnaric dialects cited above, and it implies a dissimilated */j/. This fact could suggest that the word

head', Jeh kung kuyq 'back of head'		
Sora ta?al 'spleen', Thavung kha?al 'belly, stomach, abdomen', Pacoh acheal 'heart'	*ta?al, *ti?al	PMP *tian 'belly', PC *tiān 'abdomen'
None		PMP *daRaq, PC *darah 'blood'
Santali (V7) jaŋ, Lamet sŋ?aŋ, Jehai jə?iŋ 'bone'	*cinqaŋ, *canqaŋ[41]	PMP *tuqela(n,ŋ), PC *tulāŋ 'bone'
Bonda da?tu kuĭ, PW *tɨs 'breast', Semai ntoh 'chest'	*n[s]uq[42]	PMP *susu, PC *tisəu 'breast'
Rengao ĭk krĭng, Chrau ăch câng 'ear wax', Sre klìŋ tòr 'outside part of ear'	**(n)[qḷ]eŋ[43]	PMP *taliŋa, PC *tariŋã 'ear'
Kharia (V250) mə'ḍ, PVM *mat 'eye', Proto-Plang *hàk[1] kita?[1] 'eyebrow'	*mə(n)ta(q)[44]	PMP, PC *mata 'eye'
Semai məñcɔ̀k 'fat (grease)', CN məyəć 'coconut oil'	**mə[ñc]ək(i)[45]	PMP *m(i,e)ñak 'fat/oil', PC *mañak 'oil'
Bonda tupu-sər 'oil made from marrow', surle?mi?, OM sinmor 'nasal mucus'	*saγ, *suγ[46]	PMP *himaR 'fat/oil'
Che' Wong ?ac, CN aic 'belly', Santali (K282) laĵ 'stomach, belly', Chrau talaq 'intestines'	*[?]aqi, *laqi	PMP *tinaqi 'guts'

---

was borrowed from AN, but the only possible Chamic reflex detected is Chru {khùt} 'back of head', with an anomalous earlier */gh-/ and a final */t/. The Boriwen form indicates that the final evolved more likely as */di/ > */ji/ > */j/, a regular change in AA.

41. The Austric root was apparently */qaŋ/, whence */tuqəḷaŋ/ > PAN */tuqela[n,N,ŋ]/. Such FO reflexes as Siraya */tu[q]ḷal/ suggest that the root may have originally been */qal/, but the final lateral in those forms may be secondary.

42. */s/ > /t/ in the */ns/ cluster appears to be a post-PAA development.

43. Replaces */[qə](n)[qr]iŋ/ cited in Hayes 1992:159. This comparison is somewhat speculative because it is based on the assumption that Rengao {krĭng}, Chrau {câng}, and Sre {klìŋ} meant originally 'ear'. Rengao {ĭk} and Chrau {ăch} signify 'defecate', Sre {tòr} 'ear', but the meaning of {krĭng}, {câng}, and {klìŋ} is not specified in the available glossaries. Note, however, the similar Khasi construction {eit-kor} 'ear wax', where {eit} means 'excrement' and {kor} 'ear'. Also cf. Jeh {kŭng klik} 'be deaf' and {jĭ ngĭng} 'deaf and dumb' where {kŭng} and {ngĭng} may be reflexes of the Austric root */[q]eŋ/, whence */[q][l,r]aḷeŋa/ > AN */taliŋa/, cf. AT */(N)qraḷiŋ[a]/ 'ear'.

44. Stress shift is indicated in Austroasiatic, as in */matáq/ > */mətáq/ > */məta?/ > */mət/. In a few reflexes like Proto-Plang */hàk[1] kita?[1]/ 'eyebrow' (*/hàk[1]/ 'hair'), stress shift is not apparent.

45. Apparently limited to Central Aslian and Nicobar, this word cannot be accounted for as an AN loan due to the medial cluster. Also cf. Bonda {su?} 'oil, ink' and VN {nhựa} 'sap, gum, tar, resin', which suggest the variant */(ñ)c[ə]q/.

46. This comparison is based on the assumption that the Austric root was */sVγ/ and PMP */himaR/ reflects an infixed derivative */simaγ/ of the */saγ/ root variant.

Bonda kirim-ti 'fist', Brou ndeim atêi, Khmer mrām 'finger'	*(n)ļem[a][47]	PMP *(qa)lima 'hand', PC *limã/lumã 'five'
Bahnar (PB) kŏ̆l, Jeh kãl, Mal klɨq 'head'	**(n)qoļu(q)	PMP *qulu 'head'
Kharia (V216) sɔʔlui, Mal nsook, PVM *usuk 'hair'	*(n)suk[48]	PMP *buhek 'head hair', PC *ʔbūk 'hair'
None		PMP *qaqay 'leg/foot', PC *pha 'thigh'[49]
Mon pli 'spleen', PVM *pleʔ, Sengoi pele 'fruit'	**p[a]le(q)[50]	PMP *qatay, PC *hatai 'liver'
Sora sissid, VN (*ñśic >) thịt 'flesh, meat', PM *sac 'fruit, nut, berry, acorn, pod'	*c(i,ə)ci	PMP *hesi/isi 'meat/flesh'
Pacoh piaq 'mouth, opening, end of river', Mah Meri pak, Sengoi mpak 'mouth'	**(m)paq[51]	PMP *baqbaq, PC *mubah/ bubah 'mouth'
MUK kel 'neck', Pacoh cŏ̆l 'wear around neck', Sengoi kelkeil 'ankle, wrist'	*(n)qe[R]	PMP *liqeR 'neck'
None		PMP *ijuŋ, PC *ʔ(i,u)dŭŋ 'nose'
Kharia tɑrɑn, Theng blah 'shoulder', Khasi ta-bla 'shoulder piece of animal'	*(ba)Ra(q)	PMP *qabaRa, PC *bara 'shoulder'
NK təklɛt 'scales, scab'	{**klet}[52]	PMP *kulit, PC *kulīt 'skin'
Bonda leʔaŋ, PVM *laːs, Brou liaih 'tongue'	*(di)ļaq(i)[53]	PMP *dilaq, PC *dilah 'tongue'

47. The AN numeral 'five' is derived from the word for 'hand', cf. AN (Lopez, no date) */limah/ 'five, hand'. The same derivation is observable in AA, cf. Stieng {pram} 'five' from */pļem[a]/, VN {lăm, nhăm} 'five' from */(n)ļem[a]/, and Alak {dam} 'five' from */[pa]nļem[a]/.

48. PMP */buhek/ represents perhaps a morphonemically shifted variant or doublet, cf. (*/busuk/ >) AN (Lopez) */buhuk/ 'hair', which PC */ʔbūk/ reflects instead of */buhek/.

49. It is unclear to the writer whether or not PC */pha/, which is a continuation of AN (Dempwolff 1938) */paha'/ 'leg, stalk, stem, thigh', cf. also AN (Dyen and McFarland 1970) */páqa[ʔ,h]/ 'thigh', and PC */takai/ 'foot, leg', the AN antecedent of which is unknown, are related to PMP */qaqay/. Possible AA correspondents exist for */páqa[ʔ,h]/, but not for */qaqay/.

50. By itself, this word denotes infrequently a body part, cf. also Katu {palâi} 'stomach' and Khasi {khyllái} 'the kidneys'; its more common meaning is 'fruit', but it often occurs as a classifier of such parts, cf. Chrau {play nuih} 'heart' and Rengao {pli hoʼnih} 'muscle', among others.

51. The Aslian forms may reflect */mpaːŋ/, since */ŋ/ > /k/ is common in that area, but */(m)paŋ/ probably developed from */(m)panq/ or */(m)paqpaq/.

52. The sole example, but it is difficult to account for this word by borrowing because Nyah Kur is located in Thailand and a good distance from the nearest Chamic speakers.

53. Also cf. Khasi {jliah} [jliəʔ] 'lick', which reflects most faithfully the form and possibly the meaning of PAA */di[l]aq/.

Mon daṅāp 'back tooth, molar', Sengoi lemuin tengep 'molar teeth'	**[taN]Gep[54]	PMP *ipen/nipen 'tooth'

Additional:

Sora (K115) kəmpuŋ 'belly, stomach, abdomen', Brou púng, Chrau lapong 'stomach'	*kə(m)puŋ	PEMP *kampuŋ 'belly'[55]
Kharia (V51) jurɑ?, PM *jrlaa?, Sengoi jerla 'thorn'	*ju[R]aq[56]	PCEMP *zuRi 'bone'
Pacoh arŏq, Ruc kănʒuọŋ, KY trooŋ 'throat'	*Ru(n)qaŋ	POC *Ruqa 'neck'[57]

3.6. Kinship (9/9)

Santali (V205) hən 'son, child', PM *kɔɔn 'child, offspring, young (animals)', Mintil ?awą? 'child'	*(n)qu?an[ak][58]	PMP *anak, PC ?anāk 'child'
Santali mama 'maternal uncle', Katu ama, Pacoh a-ám 'father', Bahnar ma 'younger brother of father or mother'	*(qa)ma(ma)	PMP *t-ama, PC *?amã 'father'
Pacoh a-i 'mother', Kharia (K349) bui 'girl', Mon imbay 'elder brother's wife, husband's elder sister'	*a[x]i, *bu[x]i, *mpa[x]i	PMP *bahi 'female/woman'[59]
Kharia (K535) sou 'husband', Katu sasaau 'father's cousins, sister's husband, father's sister's children', Proto-Semai *bnsaaw 'wife's elder brother'	*saw[a]	PMP *qasawa 'husband, wife'

54. Limited to Mon and Sengoi, but retention of the medial velar indicates that this is not an AN loan. Note that some Austronesianists also propose reconstruction of a velar, cf. AN (Capell 1943) */(ŋ)ipən/, (Benedict 1976) */(N)Gipən/ 'tooth'.

55. Cf. AT */[kə](m)puŋ/, AN (Dempwolff) */kə(m)puŋ/ 'belly'.

56. This comparison suggests that the postvelar spirant */R/ may represent an ancient */rl/ cluster in some cases. Also cf. AN */[d,D]uRih/ 'thorn'.

57. The POC form is probably a reflex of AN */Ruqaŋ/ 'hollow space, throat', cf. also Malay {kĕrongkong} and Chru {rɔkòng} 'throat', the latter probably a MK loan.

58. The Austric root may have been */qu?ak/, cf. Katu {vok} 'boy, son', in which case the AA and AN reconstructions are infixed forms. Mintil plausibly reflects the */[ak]/ final via */qu?an[ak]/ > */?uwaŋk/ > {?awą?} [ą = nasalized vowel], cf. also Kensiu {wɔŋ} 'child'.

59. Cf. PC */binãi/ 'virgin', AN */binay/ 'wife, woman/female'.

Thavung nAA[2] 'mother's younger brothers', Sengoi menah 'parent's younger brother', Bahnar nă 'parent's elder sibling'	**na(q)[60]	PMP *bana 'husband'
Bonda laĭbu? 'male pig', Pacoh alay, Stieng clay 'brother-in-law'	*(n)qalay	PMP *ma-Ruqanay 'male/man'
Kharia (V185) gəmkɛ 'master', Khmer ge 'one, people, someone', Bateg Nong lə? 'person'	*(lan)qe(q)[61]	PMP *laki 'male/man', PC *lakəi 'man, person'
Kharia nana 'elder sister', Bonda tuna 'younger sister (addressed by a brother), wife's younger brother's wife', Sedang na 'older sister, cousin'	*(na)na	PMP *t-ina 'mother'
Kharia lebu 'man, person', PMN *klô 'male', Semnam kəlo? 'elder sibling'	*(kal)əwu(q)[62]	PMP *tau/taumataq 'person', PC *matəu/patəu 'child-in-law'

3.7. Cultural artifacts (6/6)

Mundari (V200) oṛa? 'house, building', Bonda ŋgɔm 'village', VN (*śu:m >) thum 'hut on stilts used by hunters'	*[u]ɣaq, *(sun)ɣum[aq][64]	PMP *Rumaq 'house'[63]
Kurku (V279) jimu, jumu, OM yamo'/yimo'/himo' 'name', Chrau tanhya 'to name'	*(n)jam[u]s[65]	PMP *ŋajan, PC *ʔanə̃n/ʔaŋə̃n 'name'
VN (*[k]ləm >) trâm 'hairpin'	{**Rom}[66]	PMP *zaRum, PC *jurŭm 'needle'

60. The only known AA examples. Despite the homophony, this etymology is apparently distinct from PMP */t-ina/ 'mother'.

61. The Kharia word is apparently analyzable into {gɔm} 'house' and {kɛ} 'man', cf. Bonda {ŋgɔm} 'village' (see 'house' under Subsection 3.7).

62. The Kharia form is an isolate in Munda, but this etymology is well represented in Mon-Khmer. Pacoh {avur-avang} 'small boy' and {avŏq} 'grandfather' may reflect the root*/əwu(q)/, whence */təwu[q]/ > AN */[t]awu[h]/, but most reflexes evidence prefixation by */l/ or */k(a)l/.

63. No reflexes found in Chamic, but cf. Malay {rumah} 'house'.

64. The old */sr/ cluster, which has coalesced in Vietnamese, cf. also Katu {sum} 'shelter', is maintained in a MK loan in NRoglai {craq srum} 'hunt from a blind'. Reflexes of the root also occur in MK, cf. Pacoh {croq} 'shed, shelter, pen, sty (for pigs or buffalo)', Mon {pra'} [prɔ?] 'lean-to at end of house, used for storing paddy'.

65. The AA data suggest an ancient compound composed of */ja/ and */m[u]s/.

66. This word has been borrowed into Mon-Khmer from Chamic and Malay, cf. Katu {jarum} 'needle', but the VN form does not appear to be a loanword.

Kurku kora 'way', PW *kra?, Pear khra 'road'	*k(a,u)l̥a[67]	PMP *zalan, PC *jalān 'road'
Sora tɑre:l-ən 'string, thread', Thavung atal 'cord, thread', Pacoh ntár 'cord, single handle, basket straps'	*tal̥	PMP *talih 'rope', PC *taləi 'cord'
Khasi tap 'to cover', Mundari (V3) dɑ'b 'cover a roof, thatch', Palaung dǎp 'to cover, thatch'	*(n)təp	PMP *qatep 'thatch'

3.8. Descriptives (35/23)

Boriwen kəya? 'bad', Ta-Oy ?iji? 'false', VN dữ 'fierce, ferocious, wicked'	**ñjaq, **ñjəq[68]	PMP *zaqat 'bad', PC *jhāt 'wicked, bad'
PW *ra? 'big', Theng ya? 'far', Sora (V40) sɑŋɑ:j-ən 'be at a distance'	*(can)ɣaya(q)[69]	PMP *ma-Raya 'big'
Katu tam 'black', PM *prtam 'night (time)', Sora süm 'grope in the dark'	*(i)təm	PMP *ma-qitem 'black'
Kharia (K208) rɑŋgɑ, VN lạnh 'cold', Khmer sreň 'to cool'	*(z)l̥eŋ[70]	PMP *ma-diŋdiŋ 'cold'
Sengoi bor 'good, fine, beautiful', bernor 'goodness, righteousness, true', Pacoh nnôr 'happy'	**bər, **bənər[71]	PMP *ma-bener 'correct/true'
Mon mih 'body dirt', Mintil kamah 'dirty'	**məz[72]	PMP *cemeD (?) 'dirty'

67. This comparison is based on the assumption that the Austric root */l̥a/ is incorporated in PMP */zalan/. Also cf. Khasi {lâd} [la:t] and Pacoh {carna} 'road', the latter hypothetically from */kal̥an/ or */kal̥ana/ > */kərna:/.

68. Chamic appears to have two sets of reflexes: Rade {jhat} 'damaged, bad' clearly continues PC */jhāt/ (more likely */jahat/), but WCham {kadjah} 'bad' must be borrowed from Mon-Khmer. The MK forms suggest that the Austric root was */zVq/, whence */dzaqat/ > AN */zaqat/ and */dz(a,ə)q/ > AA */nj(a,ə)q/ > */ñj(a,ə)q/. Chrau {dagat} 'deceive' apparently reflects the variant */nzanqat/.

69. This etymology is well represented pan-Austroasiatic. Some forms suggest the variants */canqaya/, cf. Jehai {cəkey} 'big', and */Gaya/, cf. Nicobar (Car) {høi} 'far'.

70. This comparison is based on the assumption that Austric */zl̥eŋ/ > */źeŋ[an]/ or */źeŋ[gan]/ > AN (Dempwolff 1938) */Diŋin/ 'cold, cool, fresh', cf. AT */(q)ź[eŋ]en/ 'cold, cool'. Also cf. Katu {ngóón} 'cold season', Pacoh {tangan} 'cold (water, anything but people)', Katu {rangaai} 'cool', VN {nguội}'cool off, become cold' as plausible reflexes of */l̥eŋga(n,i)/.

71. The only available examples, but retention of the root and semantic differentiation are significant evidence that these are not AN loanwords.

72. Austric */-z/ > AN */-D/, AA */-z/, whence MK */-s/ > /-h/. This is the only AA example of final */z/ found to date. The question mark after the PMP reconstruction is Blust's.

Sora pɑruŋ 'be dried up', PMN *prăŋ, Khmer prāṅ 'dry, rainless'	*p(a,u)ɣaŋ	PMP *ma-Raŋaw 'dry'
Pacoh túl múl '(expressive) of blunt end', Bahnar tŭl 'dull, not pointed', NK thuɯul 'blunt, not pointed'	**tu[l][73]	PMP *pundul 'dull/blunt'
None		PMP *dumpul, PC *tump[u]l 'dull/blunt'[74]
Sora (V396a) jəru:-n, Khmu' jru? 'deep', Temiar jəro? 'long'	*jar[?]uq	PMP *zauq 'far'[75]
VN ước 'to desire, wish for, hope for', Khasi kwah [kwa?] 'wish for', Jeh wă [wa?] 'want, like, be fond of, desire'	**i[?]ak, **u[?]aq	PMP *ma-pia 'good'[76]
None		PMP *mataq 'green', PC *matah 'blue, raw'
Cua parêq, Chrau gât, Mendriq (Pang. Gal., H68) hĕnjut 'heavy'	**baRe(n)qət	PMP *ma-beReqat 'heavy'[77]
None		PMP *ma-qudip 'living/alive', PC *hadip 'to live'
None		PMP *anaduq 'long (objects)'
Kharia (V93) japi'ḍ 'shut the eyes', VN chật 'be narrow, (of clothes) be tight', Pearic čit 'tight, be close'	*(n)kit[78]	PMP *kepit 'narrow'[79]
PW *cro? 'new', Bahnar chrêu 'strange'	**qi[R]u(q)[80]	PMP *baqeRu, PC *bahrəu 'new'

---

73. Also cf. Sora {t?u:l} 'be stout' as a possible Munda correspondent. The irregular */t/:*/d/ correspondence is plausibly explained by Austric */nt/ > AN */d/.

74. PMP */dumpul/ should probably be regarded as a metathesized variant of PMP */pundul/ cited just above. The Chamic proto-form is by the writer and based on Chru {buɔl} 'dull', NRoglai {avɔn}, Rade {băl}, and WCham {tabêl} 'blunt, dull'. This word has been borrowed by Bahnaric and Katuic, cf. Bahnar {bŏl} 'dull (knife)' and Pacoh {pŏl} 'dull because of use'.

75. Cf. Chru {jruh} 'long' and NRoglai {juaq} 'length of foot', which are not clearly continuations of PMP */zauq/. Chru {jruh} is more likely a Bahnaric loan.

76. Cf. Rade {jăk} 'good', AT */[p]iya[ak]/ 'desire, hungry'. The Rade form is probably borrowed from Mon-Khmer. Also cf. Chrau {yăh} 'good' as a reflex of a variant */i[?]aq/ or */i[?]aG/.

77. PC */tr[o]?/, based on Chru {tră}, Haroi {trauq}, NRoglai {tràq}, Rade {ktrŏ}, WCham {trauk} 'heavy', is not unambiguously a continuation of PMP */ma-beReqat/.

78. Early palatalization is indicated, i.e. */(n)kit/ > */cit/ and */ñjit/, whence pre-Munda */japit/.

79. Cf. PC */kapit/ 'to close, pinch, press' from AN */ka(m)pit/ 'compress, hold together', which is related to, but not a direct reflex of PMP */kepit/.

80. Only two correspondents have been found, but they are from different branches of Mon-Khmer and non-suspect as loans.

Khmer cāʼs, Pearic čhuːs 'old', Katu takóh 'grown'	**ti(n)qas(i)[81]	PMP *ma-tuqah 'old (people)', PC *tuha 'old'
Bonda simeṛakʼ, PW *s(ŋ)krak, Bahnar 'brê 'red'	*ɣa(k,q)	PMP *ma-iRaq, PC *mahirah 'red'
Bonda səsə 'be rotten', Sora (K193a) sʔoː 'to stink', Katu nsoq 'strong fish smell, decayed, spoiled'	*su(q)	PMP *busuk 'rotten'
Kharia lorog 'to rot, decay', VN rục 'be rotten', NK phròok 'spoiled'	*ɣok	PMP *buRuk, PC *brŭk 'rotten'
None		PMP *ma-tazim, *ma-tazem 'sharp'
None		PMP *ma-babaq 'short (height)'
Bonda gĭakʼ 'shame', PW *[gac] 'ashamed, shy', Stieng haas 'feel ashamed, bashful'	*g(i,a)haq(i)	PMP *ma-hiaq 'shy/ashamed'
None		PMP *ma-sakit 'sick/painful'
None		PMP *kedi 'small'
Sora okij 'a little more', Katu machưưiq 'short time', Pearic keːč 'small'	*(n)zekiq	PMP *dikiq 'small', PC *tiki? 'small amount'
Semang (Sem. Pa. Max., H67) takăl, Mundari (Hasada)(V332) hambal 'heavy', Mal mpal 'thick'	*(n)qa[l], *qampa[l]	PMP *ma-kapal, PC *kapāl 'thick'
None		PMP *ma-nipis 'thin (materials)', PC *lupih 'thin'
None		PMP *ma-panas 'warm/hot'
Pearic pəčaʼk, Chrau suh, Tampuan tšətšuih 'wet'	**[b]acaq, **[u]c[a]q(i)	PMP *ma-baseq, PC *pasah 'wet'
Bahnar bak 'white, clean', VN bạc 'silver, money', Katu taraak 'white earring'	**bak, **[ba]rak[82]	PMP *burak 'white'
Khasi balih, Tareng blai 'white', Pacoh aturah 'white-spotted skin'	**(m)pa[l]eq[83]	PMP *ma-putiq, PC *pitih 'white'

81. The phonological development is obscure. Presumably, palatalization occurred, i.e. */tiqasi/ > */ticaś/, after which the first syllable vowel was replaced and vocalic assimilation and/or shift occurred, as in */tacaś/ > */(tə)caaś/ and */tucaś/ > */(tə)cuuś/. The first syllable of Khasi {shado} 'old' may reflect a truncated */caː/. Such forms as Rengao {kră} [kra?] 'old' and Sengoi {ngra} 'old (of people and animals)' indicate an alternate evolution, i.e. */[ta]nqra(q)/ > */ŋkraa(?)/.

82. This comparison is based on the assumption that PMP */burak/ reflects an earlier infixed form, cf. also PC */pirăk/ 'silver, money', which is probably not a direct reflex of PMP */burak/. A */buk/ variant also occurs in Mon-Khmer.

83. This comparison is based on the assumption that Austric */(pu)p[l][e]q/ > AN */putih/, cf. AT

None		PMP *ma-labeR 'wide'
None		PMP *ma-kunij, PC *k[]ñi? 'yellow'

3.9. Verbs (73/53)

Jeh kɔkùat 'detest, hate', Khmer kot 'hold in awe', Mon takuit 'take fright'	**(n)k[o]t	PMP *ma-takut 'afraid'
Sora (V334) gɑd 'cut', rɑj 'cut into small pieces as wood', PM *rac 'cut with a sickle, reap', Katu karóóch 'cut kernels off'	*(n)ɣat(i)[84]	PMP *kaRat 'bite'
Santali (V202) hɔe 'wind, air, climate, to blow', Kharia (V202) kɔyɔ 'wind', VN hiu hiu '(of wind) blow lightly'	*[q]uyu[85]	PMP *hiup 'blow', PC *?aiup 'blow, whistle'
Mundari (K537) rowa, Sengoi ruai 'soul, spirit', Pacoh rvai 'soul'	*r(a,u)wa(i)[86]	PMP *ma-ñawa 'breathe', PC *laua 'air'
Katu pa-óh 'cook', Khmer 'us 'firewood', Bahnar tɔnuh 'hearth'	**[?]us(i), **t[u]nus[87]	PMP *tunu 'burn'
Khasi pli 'change', MUK pál, pánh 'sell', Kharia paṭay 'fix price, bargain'	*pə[l̥]i[88]	PMP *beli, PC *bəli 'buy'
Sora mɑj 'to taste, lick', Khmu' mah 'eat', Katu maq 'prechew food'	*maq(i)	PMP *mamaq, PC *mumãh 'chew'
PW *ras, Pacoh rôih, Semai	**(u)l̥aqi	PMP *piliq 'choose'[89]

---

*/p[u]p[r,l]i[q]/ 'white, silver'. Benedict also reconstructs the variant AT */p[u]p[r,l]u[q]/, to which Pearic {prus} 'white' is apparently correspondent.

84. Also cf. Kharia (V334) {khɛ'ḍ} 'to bite', with {kh-} possibly reflecting earlier */kVg-/. The AA forms may reflect */ɣat/ and */ɣət/; ablauted reflexes of Austric */ɣVt/ which also underlies AN */kaRat/ 'bite, gnaw, nibble' and */keRet/ 'cut (off)'.

85. The final /0/:*/p/ correspondence is seen in other comparisons, and the labial stop is sometimes present in Austroasiatic and not in Austronesian, the reverse of the case here. It appears to have been a suffix used to form verbal and descriptive attributives from nominal roots.

86. It is assumed that */n[r]/ > */ñ/ in Austronesian, but it is possible that different initials are involved, */rV/ in Austroasiatic, */ny/ in Austronesian.

87. This comparison is based on the assumption that PMP */tunu/ reflects an affixed derivative */tunus/ of the Austric root */[?]us/.

88. The original meaning was probably 'change, exchange', cf. AT */paḷi, (m)baḷi/ 'exchange, change, buy, sell'.

89. Cf. PC */ruah/ 'choose', which is probably borrowed from Mon-Khmer and quite likely a Katuic language, cf. Brou {ruɔh} 'choose'.

(Serau, C120A) chenlas 'choose'		
None		PMP *pa-nahik 'climb'
Pearic khəčay 'climb', Khmer chkay 'stand straight up or out', VN ngay 'be straight, erect'	**ca(n)ka(i)	PMP *sakay 'climb'
PM *?a(a)r 'go, walk', Bonda ọre 'go together', Khmu' ar 'be at, stay at, take away'	*[?]aɣ	PMP *um-aRi, PC *marai 'come'[90]
None		PMP *zakan 'cook'
None		PMP *tanek, PC *tanə̃ʔ 'cook'
None		PMP *iap 'count', PC *iāp 'calculate'
None		PMP *bilaŋ 'count'
None		PMP *taŋis 'cry'
Rengao chră 'split, divide, crack open', Sre trac 'shave' Pacoh trĕq 'chop'	**taɣaq, **taɣak	PMP *taRaq 'cut (wood)', PC *tarah[91]
Katu ntaak 'chop', Pacoh tích 'chop firewood', Rengao kɑtĕk 'snap, break, cut skin'	**tak, **tek	PMP *tektek 'cut (wood)', PC *taʔ 'cut'
None		PMP *matay 'die', PC *matai 'dead'
OM kīr, Stieng kur 'dig', Sora gɑr 'throw up with shovel'	*(n)ka[l]	PMP *kali, PC *kaləi 'dig'
Chrau vĭq, Katu bâch 'lie down, sleep', Sora mimiḍ 'sleepy'	*(m)pe(qi)	PMP *hipi, PC *lupəi 'dream'
Katu ôm 'to drink', Khmu' ʔom 'water', Santali (K542) um 'bathe'	*[?]om[92]	PMP *inum, PC *minum 'drink'
Sora g(?)ɑ: 'to drink, "eat" any liquid food', Rengao kă [kaʔ], PM *caaʔ 'eat'	*(in)ka(q)[93]	PMP *kaen 'eat'
Mundari (K149b) tɔmbɔʔ 'fall forwards', Katu tampoh 'drop',	*t(a,u)(m)puq	PMP *nabuq, PC *labuh 'fall'

90. The Chamic proto-form is by the writer and based on Chru {mɔrai} and WCham {mai} 'come'.

91. Lee 1966:217 bases this PC reconstruction on Jorai {trah} 'to sculpture' and Roglai {taiah} 'trim', but cites no definition for it.

92. This comparison is based on the assumption that PMP */hinum/ is an affixed derivative of the Austric root */[?]om/ 'water', cf. AN */[d,D]anum/, AT */()(n)zalom/ 'water'. In Austroasiatic, the sense 'drink' is rare, but does occur, cf. Katu {ôm} 'to drink'. Chrau {nôm} 'urinate' evidences the same type of infixation seen in PMP.

93. Alternatively, AA */(in)ka[q,?]ən/ is reconstructible, cf. VN {ăn} 'eat', Bahnar {raqən} 'bite', Pacoh {pancan} [pənka:n] 'chew', but it is unclear whether */[q,?]ən/ is a suffixal construction or a root in compound with another root */ka/.

Bahnar puh 'slip, fall into a hole'		
Chrau hor, Khmer hūr 'flow', Pacoh lur daq 'spilled because too full, flood'	**qoɣ, **[qa]loɣ	PMP *aliR/aluR/saliR 'flow'
None		PMP *Rebek 'to fly'
Kharia (V286) mu? 'come out', MUK mọc 'grow, come up', OM mok 'appear'	*m[b]uk	PMP *tu(m)buq, PC *tumũh 'grow'
Bonda (*zəŋ >) ɔŋ, PW *hŋet 'hear', VN nghe 'listen to, hear'	*zəŋ[əɣ(i)]	PMP *deŋeR 'hear'
PMN *pôn, NKhmer pon, Pearic po:n 'hide'	**p[o]n[94]	PMP *buni 'hide'
None		PMP *palu 'hit/beat'
Kensiu cəkam, VN (*gəm >) cầm 'hold', Sora kum-si: 'hold one's fist, hold a handful'	*(n)kəm	PMP *gemgem 'hold (in fist)'
Jeh rŭp 'catch, seize', Khasi kynrup 'pounce upon, seize', Bonda sɔp' 'hold, catch', Stieng choop 'hunt'	*(n)[r]op, *(c,s)[r]op[95]	PMP *qanup 'hunt'
None		PMP *bunuq 'kill'
Khasi kynto, kyntu 'urge, persuade', Jeh pơtou 'show, guide', Chrau pađau 'teach'	**taqu[96]	PMP *taqu, PC *thəu 'know'
Bonda ḍɔḍɔ, PNB *(q)do, Kurku (V302) landa 'laugh'	*ntaw[97]	PMP *tawa 'laugh'
Nha Heun plĭp, Thavung kñiip, Bahnar 'nhĭp 'close eyes'	**(i)ləp	PMP *qinep 'lie down'
Khmer pœk, PM *pɔk 'open', Jeh pŏk chŏk 'remove, take off'	**pok	PMP *buka 'to open'[98]

---

94. David D. Thomas is gratefully acknowledged for pointing out this new AA/AN comparison, as well as two others, viz. PMP */taqu/ 'know' and */tudaq/ 'throw'. Also cf. Bahnar {ôn} 'hide something', which indicates a monosyllabic root. Possible variants, Pacoh {pul} 'to cover, hide, shelter' and Bahnar {pur} 'hide, conceal', suggest */p[o]l/, whence */p[o]li/ > */p[o]ñi/ > */po:n/, cf. also AN (Dyen-McFarland 1970) */buñi[q?h]/ 'hide'.

95. This comparison is based on the assumption that */nr/ > AN */ñ/ > PMP */n/, cf. AN (Wolff 1993) */qañúp/ 'hunt' (also see footnote 86). In Austroasiatic, cf. PMN /ñŭp/ 'grasp'.

96. The AA forms may derive from an old causative, i.e. 'cause to know' > 'teach' and by trope 'show' and 'persuade'. Also cf. Mon {tho} 'to train, employ in a trade' and PM */[d]ndoo?/ '(Nyah Kur) headman, (Mon) teach, instruct'. Reflexes of a variant, */(n)tequ/ or */(n)taqi/, are also present, cf. VN {dạy} 'teach', Khmer {phtē} 'ordain, instruct', Sengoi {panei} 'know'.

97. Limited to Munda and North Bahnaric; however, cf. VN trò 'game, trick, feat' and trêu 'tease, pester', which suggest the variant *tlaw, viz. also AT */(n)tlaw[a]/ 'laugh'.

98. PC */pŏk/ 'to open' is probably a MK loan.

Alak ntăm, Khmer ṭāṃ, OM ptaṁ 'to plant'	**təm[99]	PMP *tanem 'to plant'
None		PMP *bayu, PC *ʔiəu 'pound'[100]
Khmer ṭuk 'beat, pound', PM *kndɔk 'pound (earth)', MUK (*duk >) tục 'to chisel'	**ntuk	PMP *tuktuk, PC *tŭk 'pound'[101]
None		PMP *tutur 'say'
Sora kɑrkɑ: 'cry (as birds)', Pacoh ticár 'to crow', Khasi kren 'speak'	*(n)qaɣ(i)[102]	PMP *kaRi 'say'
None		PMP *kaRaw 'scratch (an itch)'
Pacoh lây, Chrau sây, VN thấy 'see'	**(s)[R]ai[103]	PMP *kita 'see'
Pacoh ĕh, Katu jih, Sengoi ceik 'sew'	**(n)qiq[104]	PMP *tahiq 'sew'
Katu jââch 'weave by machine', VN (*śət >) thắt 'tie, make a knot', Bahnar săt 'knit'	**jəqit[105]	PMP *zaqit, PC *jhīt 'sew'
Santali (V6) a?, PW *?ak 'bow', PM *tŋaa? '(cross)bow'	*[?]aq[106]	PMP *panaq 'shoot (arrow)', PC *panăh 'shoot'
Katu ndươl 'sit on ground'	{**d[a]l}	PMP *ma-tudan 'sit'
Pacoh chur 'sleepy or sad	*(n)zoɣ	PMP *tiduR/tuduR 'sleep'

---

99. This comparison is based on the assumption that PMP */tanem/ is a reflex of an infixed derivative */tanəm/ of the Austric root */təm/.

100. PC */ʔiəu/ is based by the writer on a single example, Chru {iơŭ} 'to strike'. It is unclear whether the Chru final glottal reflects the initial glottal transposed or an old final oral stop.

101. PC */tŭk/ is based by the writer on Chru {tơŭ}, NRoglai {tơq}, WCham {tauk} 'pound (rice)' and Rade {tŭk} 'punch'.

102. Pacoh {hear} 'to scream, squeal (pig)' indicates initial */q/, but the variant */kaɣ/ may have occurred. Also cf. PW {gray} 'to talk' and Santali {karac' kuruc'} 'mumble, grumble'.

103. This comparison is based on the assumption that the Austric root */Ra/ was prefixed differently in Austroasiatic and Austronesian. Actually, the AA forms may correspond more directly to AN */qi(n)Tay/ 'look (at)', but AN */kiTah/ apparently contains the same root, cf. AT */[ki](n)tra/ and AT */[ki]trai/ 'see'.

104. The phonetic development in Austroasiatic is obscure. EAA */(n)qis/ was cited previously, but it is possible that the earliest form was */qiq/, whence */qiq/ > */qic/ > */ʔeś/, */nqiq/ > */ñcic/ > */jiś/, and */qiq/ > */ciq/ > */ceh/ and */ceʔ/. Also cf. Kharia {cugci}, Palaung (Pan- ku) {pălɛʔ[1]} 'needle', Pearic {tey} 'sew', Tampuan {tanjia?} 'weave' as other possible members of this etymology.

105. Note stem split. Also cf. Bahnar {sĭt, chhĭt} 'sew', VN (*/siːt/ >) {tết} 'plait, braid, weave', both borrowed from Chamic, cf. NRoglai {chhiq}, Jarai {sit} 'sew'.

106. This comparison is based on the assumption that the Austric root was */[ʔ]aq/ and the AN forms are reflexes of a causative construction incorporating that root, */panaq/, which meant something like 'cause arrow or bow to shoot'. A similar causative construction is apparently present in Austroasiatic, cf. PM */pañ/ 'to shoot', but it is not clear that the AA and AN forms derive from a common Austric form. The phonology is obscure, but the AA development may have been *panáqi/ > */panáci/ > */páñci/ > */pañ/ by stress shift, syncope, and apocope. Also cf. Jehai {banəʔ} 'quiver (for darts)' and Stieng {rah} 'to shoot' as other possible examples incorporating the same root.

eyes', Birhor (V111) duṛum 'to sleep', Khmu' hmdɨr 'to snore'		
None		PMP *hajek 'smell'
Bahnar kơsoh, OM ksas 'spit', Khmer khjā'k 'spit out'	**zaq(i), **ñjaqi	PMP *luzaq 'spit'[107]
Chrau tăh 'slit open, cut up', Stieng tah 'disembowel'	**ta[q][108]	PMP *sitaq 'split'
Katu blah 'split', Kharia (V304) lα'j 'slice', Khmer -la's 'separate, detach'	*(ba)l̥aq(i)	PMP *belaq, PC *blah 'split'
None		PMP *Rames 'squeeze'
None		PMP *pereq 'squeeze'
Kurku (V381) roj(-ki) 'wring, squeeze', Chrau rêt, Sengoi rit 'squeeze'	*ɣət(s)	PMP *peRes 'squeeze'
Sora suj, VN chọc 'pierce', Sengoi cok 'stab, pierce'	*(n)cuk(i)	PMP *suksuk 'stab'
None		PMP *tuqud 'stand'
Sora ub-ye:r 'rise up', Katu (High) yir 'ascend', Thavung jool[3] 'stand up'	*(z)yeɣ	PMP *ma-diRi 'stand'
Nicobar kəlɔ:-həŋə 'steal', Mundari (V242) kumṛu, Santali kombṛo 'thief, theft, steal'	*kal̥aw, *kuml̥aw[109]	PMP *takaw 'steal'
VN tợp 'sip', Mundari (V354) si'b 'to smoke', Pacoh dyép 'suck'	*s[e]p, *(ñ)c[e]p	PMP *sepsep 'suck'
PW *ʔɐs 'swell, be swollen', Sengoi as 'swollen', Bonda buru 'make to swell'	*[ʔ]əq(i), *bur[əq]	PMP *baReq 'swell', PC *barah 'swollen'
Mundari (K519) oiyar, Ruc lọy, Riang Lang _ŋɔy 'swim'	*la(n)[ʔ]oy	PMP *laŋuy/naŋuy, PC *luai 'swim'[110]
None		PMP *demdem 'think'
Santali (V173) lebda 'throw', Khasi pda 'throw to the farthest distance possible', Chrau randăh 'throw down'	*da(q)	PMP *tudaq 'throw'
Cua takoot 'tie a knot', VN cột 'tie up, chain', Pearic khɔ:t 'tie'	**(n)kət	PMP *hiket, PC *ʔikăʔ 'tie'

107. The PMP term has been replaced by PC */kacus/ (?) 'to spit' (the question mark is Lee's), which is borrowed from Mon-Khmer.

108. Found only in these two South Bahnaric languages thus far.

109. Limited to North Munda and Nicobar.

110. It appears doubtful that PC */luai/ is a reflex of the PMP reconstruction; more likely, it is borrowed from Mon-Khmer.

Khmer ralœ̆ŋ 'uproot, overturn'	{**[l̥]eŋ}	PMP *biliŋ 'turn'
None		PMP *puter, PC *padar 'turn'
PM *taa? 'vomit', Bonda ta?mi 'sneeze', Semelai tahtɔh 'to spit'	*taq	PMP *um-utaq, PC *p_tah 'vomit'
Khasi kiew 'ascend, go up', Jeh chìu 'go, walk', PM *cəw 'return, go back'	**ikəw	PMP *lakaw 'walk/go', PC *lakāw go past, step over[111]
None		PMP *panaw 'walk/go', PC *nãu 'go'
NK khəmàa? 'dry field'	{**ma[q]}[112]	PMP *quma 'work (in garden)', PC *humã 'field'
Santali (V68) aŋgɔ'b, PM *sŋ?aap, VN ngáp 'yawn'	*(can)qap	PMP *ma-huab 'yawn'[113]

Additional:

PM *tl-?ɔɔ?, Pacoh ti-ôq, Chrau hoq 'vomit'	*u[q,?]aq	POC *luaq 'vomit'[114]

3.10. Pronouns (10/4)

Pareng (Vizagapatam) (V402) miŋgu, Khmu' ?o?, Ruc ho 'I'	*(min)qu(q)	PMP *i-aku, PC *kəu 'I'
Proto-Munda *ej 'third-person VN ấy 'that, those', Jeh i 'that'	*e[?]	PMP *si-ia 'he/she'
Bonda kɔ?na 'this one', Semaq Beri na? 'this', Ruc ná 'that'	*na(q)	PMP *i na, PC *?anãn 'that'
None		PMP *si ida 'they'
Mundari (K162) ini? 'he, this one', Khasi kane, Chrau nĕh, nìh 'this'	*(i)ne(q)	PMP *i ni/a ni, PC *?unĩ 'this'
None		PMP *kami 'we (exclusive)', PC *kamĩ 'we'
None		PMP *i-k-ita 'we (inclusive)'
None		PMP *i-kahu 'you'[115]
None		PMP *kamiu 'you (plural)'
None		PMP *i-kamu 'you (plural)'

111. PC */lakāw/ is based by the writer on Rade {hgao} and WCham {lakao} 'go past, step over'.

112. Found only in Nyah Kur, cf. AN */qumah/ 'cropland, garden, orchard'. PW */?mar/ 'field' and Chrau {mir} 'swidden field' may reflect an infixed derivative */ma[l,r]aq/, cf. AT */[(N)q]um/al̥[ah]/ 'cultivate (field), field'.

113. The only Chamic reflex found, Rade {hɔ̆ap} 'yawn', is probably borrowed from Mon-Khmer.

114. Cf. PC */?ŏ?, *h_?ŏ?/ 'to vomit', either onomatopoetic or borrowed from Mon-Khmer.

115. Cf. PC */hã/ 'you'.

None		PMP *i-k-ita 'we (inclusive)'
None		PMP *i-kahu 'you'[115]
None		PMP *kamiu 'you (plural)'
None		PMP *i-kamu 'you (plural)'

3.11. Numerals (4/2)

Kharia (K338) məiʔj, MUK mộch, PM *muay 'one'	*mu[ʔ]aca, *mu[ʔ]ica[116]	PMP *esa/isa/tasa, PC *sa 'one'
Sora ruɑ:-n, Katu madruôh 'together', Pacoh rŏi, tŏi 'pair'	*(d)ru[w]as, *(d)ru[w]ai[117]	PMP *duha, PC *dua 'two'
None		PMP *telu, PC *klәu 'three'
None		PMP *epat, PC *pāʔ 'four'

3.12. Prepositions (9/5)

Chrau avôq 'in front', Jeh bbuy 'top', Chrau vwây 'hilltop'	**bәw[q], **bab[әwi]	PMP *i babaw 'above'[118]
None		PMP *a taqas 'above'
Katu diq 'in, on', Sora (V22) tɑrɑ:ŋ-di:-n, Palaung kəndi 'middle'	*de(q)[119]	PMP *i, *di 'at'[120]
None		PMP *i babaq 'below'
PSB *pərá:m, Chrau glăm 'inside', Bahnar sơlam 'in between, in the middle'	**zaḷәm[121]	PMP *dalem 'in/inside', PC *dalăm 'inside, deep'
Thavung vɛɛl, Temoq sawɛl, Pacoh (ti-)avear 'left (side)'	**weR	PMP *ka-wiRi 'left side'
Boriwen jal jal 'near'	{**jal}[122]	PMP *hazani 'near'

116. The Austric root was apparently */tsa/, whence AA */(i,a)ca/ and then */mu[ʔ]ica, mu[ʔ]aca/ by affixation (the numeric prefix */mu/ occurs with other cardinal numbers). The latter forms were reduced by stress shift and syncope, e.g. */mu[ʔ]acá/ > */mu[ʔ]áca/ > */muac/ > MUK {mộch}.

117. Speculative. This comparison is based on the assumption that the root initial */r/ occurred optionally as prefixed */dr/, the latter shifting to AN */D/, cf. AN (Lopez) */Duwah/ 'two', whence PMP */duha/, cf. also Li */draw/.

118. Cf. AN (Dempwolff 1938) */babaw/ 'outside, surface, top (part/side)' and */ha(m)baw/ 'above/lie above, high' which apparently contain the same stem, */baw/.

119. Katu {diq} is not a Chamic loan; it is not clear that Chamic possesses a reflex of PMP */di/ (see next footnote). Also note Bonda {ḍi} 'an ablative suffix'.

120. Cf. PC */ti/, for which Lee gives no definition, based on Roglai {ti} 'in, on, at', among others. Cf. also NRoglai {di} 'from'.

121. Found only in Bahnaric thus far. The Bahnar and Chrau forms may be borrowed from Chamic, but PAA */zaḷәm/ would apparently result in Bahnar {sơlam} by regular changes.

122. Speculative. The Boriwen form is an isolate, but unlikely to have been borrowed from Austronesian, for Chamic apparently possesses no reflexes of PMP */hazani/ 'near'.

None		PMP *ka-wanan 'right side'
None		PMP *ma-taqu 'right side'

3.13. Miscellaneous (14/4)

None		PMP *amin 'all'
None		PMP *ka 'and'
VN mà 'but, yet, and, so that'	{**ma}	PMP *ma 'and'[123]
None		PMP *kuja, *kua 'how?'[124]
None		PMP *ka 'if'[125]
VN nếu 'if'	{**niw}	PMP *nu 'if'
None		PMP *qazi 'no/not'
None		PMP *diaq 'no/not'
None		PMP *liqan 'other'[126]
None		PMP *duma 'other'
None		PMP *apa 'what?'
None		PMP *p-ijan 'when?'[127]
PM *[b,p]-nəw 'where?', Thavung noo[1], Brou nau 'who'	**nəw	PMP *i nu 'where?'
VN ai, Jeh a ai 'who', Khasi ka?ey 'what'	**[?]ay or **a[?]i[129]	PMP *sai 'who?'[128]

4. Some indications of the comparison

4.1. Degree of correspondence. In Table 2, the degree of the AA/PMP lexical correspondence revealed by the basic vocabulary comparison in Section 3 is portrayed statistically. The MK/MU column gives the number of comparisons which are valid only for one or the other AA subfamily, the AA column those valid for the family. The Total column is the sum of the MK/MU and AA columns. The PMP column gives the number of PMP comparisons in each category. The AA/PMP column is the quotient of the AA column divided by the PMP column; the Tot(al)/PMP column the quotient of the Total (AA) column divided by the PMP column.

The table reveals that potential AA correspondents were found for just over two thirds (70.0%) of the PMP reconstructions. However, when only AA

123. NRoglai {mã} 'and' is the only Chamic correspondent found, and it may be a VN loan.

124. Cf. PC */kaiua/ 'because'.

125. PC */(?a)ka/ 'not yet' is a possible reflex.

126. PC */tukOn/ 'other' is not unambiguously a continuation of PMP */liqan/.

127. Cf. PC */hiã/ 'there', plausibly a reflex of AN */(q)iya[n,N]/ 'when?'.

128. No Chamic reflex has been found, but cf. Chru {aràng sơi} 'who'. WCham {hay} 'who' would appear to be borrowed from Mon-Khmer, cf. Ruc {?ay} 'who'.

129. This comparison assumes that the PAN initial is a prefix and that the Austric stem was */a[?]i/, cf. also AN */s+ai/ 'who?' where the plus sign apparently indicates that */s/ is a prefix.

Table 2. Statistical Degree of AA/PMP Correspondence

Section	Category	MK/MU	AA	Total	PMP	AA/PMP	Tot/PMP
3.2	Nature	10	12	22	29	41.4%	75.9%
3.3	Flora	4	4	8	9	44.4	88.9
3.4	Fauna	4	6	10	16	37.5	62.5
3.5	Anatomy	8	12	20	23	52.2	87.0
3.6	Kinship	1	8	9	9	88.9	100.0
3.7	Cultural artifacts	1	5	6	6	83.3	100.0
3.8	Descriptives	11	12	23	35	34.3	65.7
3.9	Verbs	23	30	53	73	41.1	72.6
3.10	Pronouns	0	4	4	10	40.0	40.0
3.11	Numerals	0	2	2	4	50.0	50.0
3.12	Prepositions	4	1	5	9	11.1	55.6
3.13	Miscellaneous	4	0	4	14	0.0	28.6
Total		70	96	166	237	40.5%	70.0%

correspondences with representative forms in both AA subfamilies are considered, the degree of correspondence falls dramatically to about two fifths (40.5%).

In Table 3, the PC correspondences are brought into the picture. It is noteworthy that the PC/PMP correspondence (43.0%) is only slightly higher than the AA/PMP correspondence (40.5%), even though PC correspondences outnumber the AA in six of the 12 categories.

Table 3. Statistical Degree of PC/PMP Lexical Correspondence

Section	Category	AA	PC	PMP	AA/PMP	PC/PMP
3.2	Nature	12	14	29	41.4%	48.3%
3.3	Flora	4	5	9	44.4	55.6
3.4	Fauna	6	8	16	37.5	50.0
3.5	Anatomy	12	16	23	52.2	69.6
3.6	Kinship	8	4	9	88.9	44.4
3.7	Cultural artifacts	5	4	6	83.3	66.7
3.8	Descriptives	12	14	35	34.3	40.0
3.9	Verbs	30	28	73	41.1	38.4
3.10	Pronouns	4	4	10	40.0	40.0
3.11	Numerals	2	4	4	50.0	100.0
3.12	Prepositions	1	1	9	11.1	11.1
3.13	Miscellaneous	0	0	14	0.0	0.0
Total		96	102	237	40.5%	43.0%

4.2. Implications of the statistics. The degree of lexical correspondence between Austroasiatic and Malayo-Polynesian indicated by the statistics cited above is surprisingly high. If the AA lexical data had come from a conventional basic vocabulary list usable in a glottochronological computation, the inferred time depth of the

PAA/PMP relationship would be less than that found by other researchers to exist between some branches of Mon-Khmer. The methodological approach used in discovering AA correspondents has obviously and significantly skewed the degree of correspondence in favor of the proposition that Austroasiatic and Austronesian are genetically related. In the writer's opinion, however, the fact that this distortion exists takes nothing away from the validity of this correspondence or its value in providing probative evidential support for the verity of that proposition.

The degree of lexical correspondence indicated by Table 2 thus implies that Austroasiatic and Austronesian are indeed genetically related and perhaps more closely than even Schmidt may have envisioned. When one recalls that loanwords have been eliminated from the comparison, no other conclusion seems possible.

5. Conclusion

The nature and degree of the lexical correspondence portrayed above lead to only one conclusion: Austroasiatic and Austronesian are genetically related. As such, they have inherited a large volume of common lexical material of both basic and non-basic vocabulary types, only some of which has been demonstrated in Section 3. When the probative phonological and morphological evidence Schmidt presented in 1906 are considered together with the lexical evidence shown here, it becomes clear that he was right after all; Austric is a valid language construct and an appropriate taxonomic entity uniting the AA and AN language families.

REFERENCES

Banker, John and Elizabeth and Mơ. 1979. *Bahnar Dictionary, Plei Bong-Mang Yang Dialect*. Huntington Beach: SIL.

Barker, Milton and Muriel. 1976. *Mường-Vietnamese-English Dictionary*. Microfiche. Huntington Beach: SIL.

Benedict, Paul K. 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots*. New Haven: HRAF Press.

Benedict, Paul K. 1976. "Austro-Thai and Austroasiatic." *AS*:1–36.

Benedict, Paul K. 1990. *Japanese/Austro-Tai*. Ann Arbor: Karoma Publishers.

Benjamin, Geoffrey. 1976a. "Austroasiatic subgroupings and prehistory in the Malay Peninsula." *AS*:37–128.

Benjamin, Geoffrey. 1976b. "An outline of Temiar grammar." *AS*:129–87.

Bhattacharya, Sudhibhushan. 1966. "Some Munda etymologies." *SICAL*:28–40.

Bhattacharya, Sudhibhushan. 1968. *A Bonda Dictionary*. Poona: Deccan College Post-Graduate and Research Institute.

Biligiri, H.S. 1965. *Kharia: Phonology, Grammar and Vocabulary*. Poona: Deccan College, Postgraduate and Research Institute.

Bishop, Nancy M. and Mary M. Peterson. 1994. "Kensiw glossary." *MKS* XXIII: 163–95.

Blood, Henry F. 1968. *A Reconstruction of Proto-Mnong.* Grand Forks: SIL-University of North Dakota.
Blust, Robert. 1993. "Central and central-eastern Malayo-Polynesian." *OL* 32.2: 241–94.
Costello, Nancy A. 1971. *Katu Vocabulary.* Saigon: Bộ Giáo Dục.
Diffloth, Gérard F. 1977. "Mon-Khmer initial palatals and 'substratumized' Austro-Thai." *MKS* VI:39–58.
Diffloth, Gérard F. 1980. "The Wa languages." *LTBA* 5.2.
Diffloth, Gérard F. 1984. *The Dvaravati-Old Mon Language and Nyah Kur. Monic Language Studies 1.* Bangkok: Chulalongkorn University.
Diffloth, Gérard F. 1990. "What happened to Austric?" *MKS* XVI-XVII:1–9.
Diffloth, Gérard F. 1994. "The lexical evidence for Austric, so far." *OL* 33.2:309–322.
Ferlus, Michel. 1974. "Lexique souei-français." *ASEMI* V(1):141–59.
Ferlus, Michel. 1979. "Lexique thavung-français." *CLAO* 5:71–94.
Ferlus, Michel. 1983. "Essai de phonétique historique du Môn." *MKS* XII:1–90.
Ferlus, Michel. 1991. "Vocalisme du proto-viet-muong." Paper circulated at the 24th International Conference on Sino-Tibetan Languages and Linguistics, Bangkok and Chiang Mai, Thailand.
Gradin, Dwight and Thong. 1979. *Jeh Vocabulary.* Huntington Beach: SIL.
Gregerson, Kenneth and Marilyn. 1977. *Rengao Vocabulary.* SIL.
Guy, Jacques B.M. 1995. "On lexicostatistics and glottochronology." *Dhumbadji! (Journal for the History of Language)* 2.1:55–87.
Haupers, Lorraine and Ralph. 1991. *Stieng-English Dictionary.* Manila: SIL.
Hayes, La Vaughn H. 1992. "On the track of Austric, part I: introduction." *MKS* XXI:143–77.
Hayes, La Vaughn H. 1997a. "Another look at final spirants in Mon-Khmer." *MKS* XXVI:51–64.
Hayes, La Vaughn H. 1997b. "On the track of Austric, part II: consonant mutation in early Austroasiatic." *MKS* XXVII:13–41.
Headley, Robert K. Jr. 1977. "A Pearic vocabulary." *MKS* VI:69–149.
Jacob, Judith M. 1974. *A Concise Cambodian-English Dictionary.* London: Oxford University Press.
Johnston, Richard. 1969. "Kuy basic word list." *MKS* III:1–4.
Kuiper, F.B.J. 1948. "Munda and Indonesian." *Orientalia Neerlandica, a Volume of Oriental Studies*:372–401. Leiden: A.W. Sijthoff.
Lee, Ernest W. 1966. "Proto-Chamic Phonologic Word and Vocabulary." Ph.D. thesis, Indiana University.
Lehmann, Winfred P. 1992. *Historical Linguistics, Third Edition.* Routledge: New York.
Lindell, Kristina. 1974. "A vocabulary of the Yuan dialect of the Kammu language." *Acta Orientalia* 36:191–207.
Miller, Carolyn and John. 1967. *English-Brou Language Familiarization Manual.* Saigon: SIL.
Miller, John D. 1994. "Evaluation of the wordlist used in a Mon-Khmer research project in Northeast Thailand." *MKS* XXIII:67–82.
Nguyễn Đình Hoà. 1971. *Vietnamese-English Student Dictionary.* Carbondale: Southern Illinois University Press.

Nguyễn Phú Phong, Trần Trí Dõi and Michel Ferlus. 1988. *Lexique Vietnamien-Ruc-Français, Parler d'une minorité ethnique des montagnes de Quang Binh, Vietnam, avec Notes ethnographiques et introduction linguistique*. Paris: Sudestasie.

Pinnow, Heinz-Jürgen. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Otto Harrassowitz.

Pinnow, Heinz-Jürgen. 1966. "A comparative study of the verb in the Munda languages." *SCAL*:96–193.

Rabel, Lili. 1961. *Khasi, a Language of Assam*. Baton Rouge: Louisiana State University.

Ramamurti, G.V. 1938. *Sora-English Dictionary*. Madras: Superintendent, Gov't Press. Reprinted 1986, Delhi: Mittal Publications.

Schmidt, Wilhelm. 1904. "Grundzüge einer lautlehre der Khasi-sprache in ihren beziehungen zu derjenigen der Mon-Khmer-sprachen. Mit einem anhang: die Palaung-, Wa- und Riang-sprachen des mittleren Salwin." *Abhandlungen der bayerischen Akademie der Wissenschaft* I.22.3:677–810.

Schmidt, Wilhelm. 1905. "Grundzüge einer lautlehre der Mon-Khmer-sprachen." *Denkschrifte der Akademie der Wissenschaft Wien*, Phil.-hist. Kl. 51:1– 233.

Schmidt, Wilhelm. 1906. "Die Mon-Khmer-völker, ein bindeglied zwischen völkern Zentralasiens und Austronesiens." *Archiv für Anthropologie*, Braunschweig, new series, 5:59–109.

Shorto, H.L. 1962. *A Dictionary of Modern Spoken Mon*. London: Oxford University Press.

Shorto, H.L. 1971. *A Dictionary of the Mon Inscriptions from the Sixth to the Sixteenth Centuries*. London: Oxford University Press.

Singh, U. Nissor. 1906. *Khasi-English Dictionary*, eds. Major P.R.T. Gurdon, U Dohory Ropmay, and U Hajom Kissor Singh. Reprinted 1966, New Delhi: Gyan Publishing House.

Skeat, Walter W. and Charles O. Blagden. 1906. *Pagan Races of the Malay Peninsula, Volume II*. London: Macmillan. Reprinted 1966, New York: Barnes & Noble.

Smalley, William A. 1961. *Outline of Khmu' Structure*. New Haven: American Oriental Society.

Smith, Kenneth D. 1972. *A Phonological Reconstruction of Proto-North-Bahnaric (Language Data, Asian-Pacific Series, No. 2)*. Santa Ana: SIL.

Theraphan L. Thongkum. 1984. *Nyah Kur (Chao Bon)-Thai-English Dictionary, Monic Language Studies 2*. Bangkok: Chulalongkorn University.

Thomas, David D. and Thổ Sáng Lục. 1966. *Chrau Vocabulary*. Saigon: Bộ Giáo Dục. Reprinted 1977, Huntington Beach: SIL.

Thomas, David and Sophana Srichampa, eds. 1996. *Phraya Prachakij-karacak, Some Languages of Siam*. Translated and annotated. Bangkok: Mahidol University.

Watson, Richard and Saundra and Cubuat. 1979. *Pacoh Dictionary*. Huntington Beach: SIL.

Wolff, John U. 1993. "The PAN phonemes *ñ and *N." *OL* 32(1):45–61.

Wurm, S.A. and B. Wilson. 1983. *English Finderlist of Reconstructions in Austronesian Languages (Post-Brandstetter), Pacific Linguistics Series C33*. Canberra: Australian National University.

2021 Biltmore Drive
Fayetteville, NC 28304
USA

e-mail: LV Hayes <lvhayes@worldnet.att.net>