

THE CURRENT STATUS OF AUSTRIC

A review and evaluation of the lexical and morphosyntactic evidence

Lawrence A. Reid

Introduction

The purpose of this chapter is to review and evaluate the set of evidence that has so far appeared in support of a genetic relationship for the Austric family of languages, here defined as constituting the AN family as its eastern branch and the AA languages as its western branch. It thereby excludes consideration of evidence which suggests that the Tai-Kadai family of languages might be included as part of the family and avoids the obfuscation that discussion of the Austro-Tai hypothesis has had on the basic question of the genetic relationship of AA and AN.

There have been a number of articles, beginning with Schmidt (1906) that have presented sets of corresponding lexical items purporting to establish a genetic relationship between AA and AN. Much of this work has been shown to be spurious, but Diffloth (1994) presents a number of what he terms 'lexical agreements' between the two families which he considers to be probable. Subsequent work by Hayes (1997, 1999) has introduced a considerable number of new equations into the arena especially in the area of so-called 'basic vocabulary' that need to be evaluated. Some of these have already appeared in earlier work, but are reintroduced to us by Hayes in his attempt to show that, although, as Diffloth (1994: 312) says, 'the lexical evidence is not impressive, it is undoubtedly there', especially in that area of the lexicon that counts most strongly towards the establishment of a genetic relationship.

Schmidt (1906) was also the first to draw attention to the striking morphological comparisons that exist between the two families. Reid (1994, 1999) expanded on this work and noted also certain syntactic characteristics, which along with the reconstructed morphology suggested an ergative structure for the parent of the two families.

The first part of the chapter will be a detailed evaluation of the basic vocabulary comparisons between PAA and PAN proposed by Hayes (1999), to determine to what degree they may be said to constitute a body of cognates, supported by

the usual requirements of recurrent sound correspondences and reasonable semantic equivalence. Hayes does not specifically claim that the pairs of forms he cites are cognates (this term does not appear in his paper), he refers to them by the less strict label of 'lexical comparison', a term which allows for forms which may be similar, not only because they are cognate, but also because they may be borrowings from one group into the other, or they may be the result of universal phonological developments, or they may simply be similar by chance. In order for the comparanda to constitute true cognates, it is imperative that a clear set of recurrent sound correspondences be established between the two proto-languages, and that the forms being compared have reasonably similar semantics.

The second part of the chapter will summarise the morphological evidence that has been proposed as evidence for a genetic relationship between AA and AN, and will discuss some of the alternative hypotheses that have been proposed to account for this evidence.

The lexical evidence for Austric

An adequate evaluation of Hayes' comparisons should consist of at least three parts: (1) an evaluation of the status of his PAA reconstructions and the methodology that he used to establish them, (2) an evaluation of the PAN reconstructions used in the comparisons and (3) an evaluation of the phonological correspondences and semantic features that supposedly relate the forms.

Hayes' PAA reconstructions

It is unclear from Hayes' paper whether the forms that he cites as evidence for each of his reconstructions constitute the total sum of his available evidence. I suspect that they probably do not, and that the few forms that he cites are representative of a (much?) larger body of evidence for which there was no room in the publication. However, I must assume that the forms that he cites constitute the best evidence available for his reconstructions.

For his PAA reconstructions, Hayes claims that most phonemic correspondences between the lexical items cited 'are in fact regular, at least where the consonants are concerned' (1999: 7), although in few cases does he attempt to make explicit what those regular correspondences are. As a non-specialist in AA languages, I have had to take this statement at face value in order to make my evaluation of his comparisons with PAN, although I suspect that a good deal of ingenuity was required in some cases to actually make the correspondences work. Hayes is, however, careful to indicate the relative time-depth of his own reconstructions.

The best claim to PAA status are those forms that he claims have reflexes in both the eastern and the western branches of the family. Of the approximately

150 forms that he reconstructs, 76 have proposed reflexes from both major branches of the family. Of the remaining reconstructions, 70 have proposed reflexes in more than one language in one or more of the EAA subfamilies. The remaining forms have reflexes in only one language, but Hayes claims that in these cases the comparisons are, in effect, too good to be ignored. I have relabeled his reconstructions that do not have a western AA reflex as PEAA, and consider that these have lower probative value than those that can justly be claimed to be PAA.

The Austronesian reconstructions

On the AN side, all of Hayes' comparisons are with reasonably well-established reconstructions. Hayes used the list of 200 basic vocabulary reconstructions for PMP, provided in Blust (1993), rather than with the smaller set that have been reconstructed for PAN. However, in evaluating the comparisons, I have chosen to compare Hayes' AA reconstructions with PAN rather than with the historically subsequent PMP, whenever an appropriate PAN form exists. The PAN reconstructions with which I made my comparison are those which are also summarised alongside the PMP reconstructions in Blust (1999). Of the 150 or so comparisons for which Hayes cites a PMP reconstruction, there are some 79 for which the PMP form is a continuation of a reconstructed PAN reconstruction. This set potentially has high probative value when compared with an AA reconstruction. The 52 forms which compare with a PAA reconstruction (marked with a single asterisk in Table 8.1) have the highest value, while the remaining 27 that compare with only a PEAA reconstruction (marked with a double asterisk in Table 8.1) are of lower value. However, given the possibility that the Munda languages may have split off prior to the split of the eastern branch from pre-AN, this set of lexical comparisons may take on greater significance. The full set of potentially comparable PAA (and PEAA) forms with presently reconstructed PAN forms is shown in Table 8.1.

It is apparent that in making his comparisons, Hayes has operated on the assumption that PA must have been a highly affixing language, with PAN inheriting many of the forms in their affixed state, while their corresponding PAA forms were inherited either as roots, or with different affixes. Table 8.2 gives a list of some of the PAN forms that have been reanalysed by Hayes as originally consisting of a root plus one or more affixes.

Table 8.3 shows many of the proposed PA affixes that Hayes implies are present in his AN reconstructions. Although some of these forms may indeed have been affixes in the putative parent of PAN (as they are in PAN and some of its daughter languages) and some of the PAN reconstructions may have been morphologically complex (e.g. *Si-kan 'fish', *ma-ka-Sepal 'thick', *C<in>aqi 'guts', even *q<al>ejaw 'day', on the basis of comparisons such as Bontok *ʔalgew* 'sun, day', reflecting the full, infixated form, with reflexes of apparently unaffixed *maʔégew* 'to be fine, after rain', *ʔagʔagew* 'morning', *maggew* 'handsome [bright appearance?]', etc.), it is methodologically unwise to equate any

Table 8.1 Proposed lexical correspondences between PAA (and PEAA) and PAN (79)

	<i>PAA and PEAA</i>	<i>PAN</i>	<i>Gloss</i>
1.0	Nature		
1.1	*qabuh	*qabu	ashes
1.2	**[j](a)raw	*qalejaw	day
1.3	*[s]uy, *[sa](m)puy(s)	*Sapuy	fire
1.4	**[əw]	*danaw	lake
1.5	**b(i,a)lal	*bulaN	moon
1.6	**ka[nj]jal	*quzaN	rain
1.7	*qa(m)puc(i)	*timus	salt
1.8	**si[y]aq, **su[y]ak	*qasiRa	salt
1.9	**[ay]	*qenay	sand
1.10	*(m)pə!	*qebel	smoke
1.11	*[o]q(i)	*bituqen	star
1.12	**tamuq	*batu	stone
1.13	*[ʔ]om	*daNum	water
2.0	Flora		
2.1	** (m)b[oʔa]q	*buaq	fruit
2.2	**k(i,a)hi(uq)	*kaSiw	wood
3.0	Fauna		
3.1	*cu(q)	*asu	dog
3.2	*teloR	*qiCeluR	egg
3.3	*(n)qa(q)	*Sikan	fish
3.4	**k[o]t(i)	*kuCu	head louse
3.5	*b[əw]	*labaw	rat
3.6	*[su][l]aR	*SulaR	snake
4.0	Anatomy		
4.1	**ko[d(i)]	*likud	back
4.2	*taʔal, *tiʔal	*tiaN	belly
4.3	*cinqan, *canqan	*CuqelaN	bone
4.4	*n[s]uq	*susu	breast
4.5	** (n)[q!eŋ]	*Calija	ear
4.6	*mə(n)ta(q)	*maCa	eye
4.7	*səy, *suŋ	*SimaR	fat/oil
4.8	*[ʔ]aqi, *laqi	*Cinaqi	guts
4.9	*(n)ləm[a]	*(qa)lima	hand
4.10	** (n)qo!u(q)	*qulu	head
4.11	**p[a]le(q)	*qaCay	liver
4.12	*c(i,ə)ci	*Sesi/isi	meat/flesh
4.13	*(n)qe[R]	*liqeR	neck
4.14	*(ba)Ra(q)	*qabaRa	shoulder
4.15	**[taN]Gep	*nipen	tooth
5.0	Kinship		
5.1	*(qa)ma(ma)	*t-ama	father
5.2	*a[x]i, *bu[x]i, *mpa[x]i	*bahi	female/woman
5.3	*(n)qalay	*ma-RuqaNay	male/man
5.4	*(na)na	*t-ina	mother
5.5	*(kal)əwu(q)	*Cau	person

(Table 8.1 continued)

Table 8.1 Continued

	<i>PAA and PEA</i>	<i>PAN</i>	<i>Gloss</i>
6.0	Cultural artifacts		
6.1	**[u]ʔaq, *(sun)ʔum[ʔaq]	*Rumaq	house
6.2	*(n)jam[u]s	*ŋajan	name
6.3	{**Rom}	*zaRum	needle
6.4	*k(a,u)ʔa	*zalan	road
6.5	*taʔ	*CaliS	rope
7.0	Descriptives		
7.1	*(can)ʔaya(q)	*ma-Raya	big
7.2	**qi[R]u(q)	*ma-baqeRu	new
7.3	**ti(n)qas(i)	*ma-tuqaS	old (people)
7.4	*ʔok	*ma-buRuk	rotten
7.5	*g(i,a)haq(i)	*ma-Siaq	shy/ashamed
7.6	*(n)qa[ʔ], *qampa[ʔ]	*ma-kaSepal	thick
8.0	Verbs		
8.1	**k[ʔ]oʔt	*ma-takut	afraid
8.2	*(n)ʔat(i)	*kaRat	bite
8.3	*[q]uyu	*Siup	blow
8.4	**[ʔ]ʔus(i), **t[ʔ]nus	*CuNuh	burn
8.5	*pə[ʔ]i	*beli	buy
8.6	**u[ʔ]aʔi	*piliq	choose
8.7	**taʔaq, **taʔyak	*taRaʔ	cut (wood)
8.8	**tak, **tek	*tektek	cut (wood)
8.9	*(n)ka[ʔ]	*kalih	dig up
8.10	*(m)pe(qi)	*Sepi	dream
8.11	*(in)ka(q)	*kaen	eat
8.12	**qoʔ, **[q]aʔloʔ	*qaluR (?)	flow
8.13	*(n)kəʔm	*gemgem	hold (in fist)
8.14	*(n)[r]oʔp, *(c,s)[r]oʔp	*qaNup	hunt
8.15	*ntaw	*Cawa	laugh
8.16	**s[R]ai	*kita	see
8.17	**nqiʔ	*taSiq	sew
8.18	*[ʔ]aʔ	*panaq	shoot
8.19	*(n)zoy	*tuduR	sleep
8.20	*(z)yey	*diRi	stand
8.21	*kaʔaw, *kumʔaw	*Cakaw	steal
8.22	*s[e]p, *(h)ʔ[e]p	*sepsep	suck
8.23	*[ʔ]əʔq(i), *bur[əʔq]	*baReq	swell
8.24	*ʔa(n)[ʔ]oy	*Naŋuy	swim
8.25	*taʔ	*utaq	vomit
8.26	{**ma[ʔ]}	*qumah	work in fields
8.27	*(can)qap	*ma-Suab	yawn

non-corresponding set of phonemes with an affix, unless justification can be found for it in the daughter languages. Such forms have therefore been eliminated from the set of potential cognates. This is not a trivial concern. Ignoring non-agreeing segments by calling them affixes without justification, allows for the inclusion of almost any non-cognate form into the comparative set.

Table 8.2 Morphological reanalysis of PAN reconstructions (implied in Hayes 2000) (26)

<i>PAN (Blust 1999)</i>	<i>Gloss</i>	<i>Hayes' Reanalysis</i>
*qalejaw	day	*q<al>ejaw
*Sapuy	fire	*S<ap>uy, *Sa-puy
*qasiRa	salt	*qa-siRa
*bituqen	star	*bituq-en
*daNum	water	*d<aN>um
*Sikan	fish	*Si-ka-n
*CuqelaN	bone	*Cu-q<el>aN
*Caliŋa	ear	*Caliŋ-a
*SimaR	fat/oil	*S<im>aR
*Cinaqi	guts	*C<in>aqi(?)
*(qa)lima	hand	*(qa)lim-a
*nipen	tooth	*nip-en
*Cau	person	*Ca-u
*Rumaq	house	*R<um>aq, *Rum-aq
*ŋajan	name	*ŋaja-n
*zalan	road	*zala-n
*CaliS	rope	*Cal-iS
*ma-kaSepal	thick	*ka-Se-pal
*Siup	blow	*Siu-p
*kalih	dig up	*kal-ih
*kaen	eat	*ka-en
*qaluR (?)	flow	*q<al>uR, *qa-luR
*Cawa	laugh	*Caw-a
*panaq	shoot	*p<an>aq
*diRi	stand	*diR-i
*baReq	swell	*b<aR>eq, *baR-eq

Table 8.3 Hayes' proposed PA affixes in PAN reconstructions

Infixes	-al-, -an-, -aN-, -ap-, -aR-, -um-, -im-, -in-, -el-
Prefixes	qa-, Si-, Cu-, ka-, Se-
Suffixes	-en, -n, -u, -iS, -a, -aq, -i

The remaining set of PAN comparable forms are given in Table 8.4.

After eliminating the proposed sets that contain a PAN form that Hayes implies is morphologically complex, there remain some 19 sets, shown in Table 8.5.

The remaining AN reconstructions that Hayes cites are PMP forms that can be grouped into three types: (1) those which continue a PAN form but which have undergone an irregular phonological change, such as metathesis (M); (2) those which constitute a lexical replacement of an earlier PAN form (L) or are an innovation alongside a PAN form than has undergone a semantic shift in PMP (I); and (3) those for which no PAN form has as yet been reconstructed (?). These forms are shown in Table 8.6.

Table 8.4 Potential PAA–PAN comparisons (31)

	<i>PAA</i>	<i>PAN</i>	<i>Gloss</i>
1.1	*qabuh	*qabu	ashes
1.7	*qa(m)puc(i)	*timus	salt
1.10	*(m)pəc	*qebel	smoke
3.1	*cu(q)	*asu	dog
3.2	*te oR	*qiCeluR	egg
3.5	*b[əw]	*labaw	rat
3.6	*[su][l]aR	*SulaR	snake
4.2	*taʔal, *tiʔal	*tiaN	belly
4.4	*n[s]uq	*susu	breast
4.6	*mə(n)ta(q)	*maCa	eye
4.12	*c(i,ə)ci	*Sesi/isi	meat/flesh
4.13	*(n)qe[R]	*liqeR	neck
4.14	*(ba)Ra(q)	*qabaRa	shoulder
5.1	*(qa)ma(ma)	*t-ama	father
5.2	*a[x]i, *bu[x]i, *mpa[x]i	*bahi	female/woman
5.3	*(n)qalay	*ma-RuqaNay	male/man
5.4	*(na)na	*t-ina	mother
7.1	*(can)ʔaya(q)	*ma-Raya	big
7.4	*yok	*ma-buRuk	rotten
7.5	*g(i,a)haq(i)	*ma-Siaq	shy/ashamed
8.2	*(n)yat(i)	*kaRat	bite
8.5	*pə[l]i	*beli	buy
8.10	*(m)pe(qi)	*Sepi	dream
8.13	*(n)kəm	*gemgem	hold (in fist)
8.14	*(n)[r]op, *(c,s)[r]op,	*qaNup	hunt
8.19	*(n)zoʔ	*tuduR	sleep
8.21	*ka aw, *kum aw	*Cakaw	steal
8.22	*s[e]p, *(ñ)c[e]p	*sepsep	suck
8.24	*la(n)[ʔ]oy	*Najuy	swim
8.25	*taq	*utaq	vomit
8.27	*(can)qap	*ma-huab	yawn

Of these groups, I have only included those marked with (?) in Table 8.6 as possible PMP comparisons, in that they may constitute a continuation of a PAN form which no longer exists in Formosan languages. These forms are considered to have lower probative value than true PAN reconstructions, and are labeled as PMP. From this set I have likewise eliminated those that Hayes implies were retentions of morphologically complex PA forms. The remaining set is provided in Table 8.7.

The final group of potentially comparable sets, and those that are of least value to supporting an Austric hypothesis are those that compare a PEAA form with a PMP form. This set (minus those that Hayes implies were retentions of morphologically complex PA forms) is provided in Table 8.8.

Sagart (email comm. 2001) has suggested ‘some of the evidence for Austric is also extra-Formosan rather than PAN, suggesting again an early contact relationship rather than a genetic one’. If in fact there was a post-PAN return to the

Table 8.5 Potential PEAA–PAN comparisons (19)

	<i>PEAA</i>	<i>PAN</i>	<i>Gloss</i>
1.5	**b(i,a)al	*bulaN	moon
1.6	**ka[ñj]al	*quzaN	rain
1.9	**lay	*qenay	sand
1.12	**tamuq	*batu	stone
2.1	**k(m)b[oʔa]q	*buaq	fruit
2.2	**k(i,a)hi(up)	*kaSiw	wood
3.4	**k[o]t(i)	*kuCu	head louse
4.1	**ko[d(i)]	*likud	back
4.10	** (n)qo[u(q)]	*qulu	head
4.11	**p[a]le(q)	*qaCay	liver
7.2	**qi[R]u(q)	*ma-baqeRu	new
7.3	**ti(n)qas(i)	*ma-tuqaS	old (people)
8.1	** (n)k[o]t	*ma-takut	afraid
8.4	**[ʔ]us(i), **t[u]nus	*CuNuh	burn
8.6	** (u)laqi	*piliq	choose
8.7	**tayaq, **taɣak	*taRaq	cut (wood)
8.8	**tak, **tek	*tektek	cut (wood)
8.2	** (s)[R]jai	*kita	see
8.2	** (n)qiq	*taSiq	sew

mainland, it is possible that some of the lexical sets proposed by Hayes which do not have a PAN reconstruction may be evidence for that. However, even these sets need to be critically evaluated, and are beyond the scope of this chapter.

The sound correspondences

My procedure was to begin with the reconstructed consonant system of PAN and PMP (Blust 1999: 34) shown in Table 8.9, and to compare, in order, each phoneme with the apparently corresponding phoneme in each of Hayes' comparisons.

It should be noted that I made no attempt to compare the vocalic systems. Although Hayes has made a provisional reconstruction of a 6-vowel system for PAA, he notes that 'additional vowel phonemes and diphthongs will probably have to be reconstructed eventually'. Comparison was based on the distributional features of the PAN phonemes in each of the (usually) disyllabic AN reconstructions. In many cases (100/154) the comparable form in Hayes was also disyllabic, and the determination of the appropriately corresponding phoneme was not difficult. The remaining forms, however, needed to be compared with what Hayes has reconstructed on the AA side with a monosyllabic form. For some 100 forms, the correspondence is between the final syllable of a PAN form (or the putative monosyllabic root of those forms that Hayes believes are continuations of affixed PA forms) and a PAA monosyllable. The full set of correspondences are shown in Table 8.10. Shading indicates distinctly different AA correspondences for the same PAN proto-phoneme.

Table 8.6 Hayes' PAA–PMP comparisons (36)

	<i>PAA</i>	<i>PMP</i>	<i>Gloss</i>
1.0	Nature		
1.1	*ɣ[a]mb[o]	*Rabun (L)	cloud
1.2	*buk	*qabuk (?)	dust
1.3	*teq	*taneq (I)	earth, soil
1.4	*(m)put	*kabut (?)	fog, mist
1.7	*(bi)lat(i)	*kilat (M)	lightning
1.9	*gər	*gurgur (L)	thunder
2.0	Flora		
2.2	*(m)puŋ	*buŋa (?)	flower
2.3	*(n)je	*baliji (L)	grass
2.5	*(s)uɣ(at)	*uRat (L)	root
2.6	*(n)qay	*wakaR (L)	root
4.0	Anatomy		
4.2	*(n)suk	*buhək (M)	head hair
4.4	*(di)laq(i)	*dilaq (I)	tongue
5.0	Kinship		
5.1	*(n)quʔan[ak]	*anak (?)	child
5.2	*saw[a]	*qasawa (?)	spouse
5.4	*(lan)qe(q)	*laki(I)	male/man
6.0	Cultural artifacts		
6.1	*(n)təp	*qatep (?)	roof/thatch
7.0	Descriptives		
7.2	*(i)təm	*ma-qitem (L)	black
7.3	*(z)leŋ	*ma-diŋdiŋ (?)	cold
7.6	*p(a,u)yaŋ	*ma-Raŋaw (L)	dry
7.8	*jar[ʔ]uq	*ma-zauq (L)	far
7.11	*(n)kit	*kepit (?)	narrow
7.12	*ɣa(k,q)	*ma-iRaŋ (L)	red
7.13	*su(q)	*ma-busuk (I)	rotten
7.14	*(n)zekiq	*dikiq (I)	small
8.1	*ɾ(a,u)wa(i)	*maŋawa (?)	breathe
8.0	Verbs		
8.2	*maq(i)	*mamaq (L)	chew
8.4	*[ʔ]ay	*maRi (?)	come
8.5	*[ʔ]om	*inum (?)	drink
8.6	*t(a,u)(m)puq	*nabuq (?)	fall
8.7	*m[b]uk	*tu(m)buq (?)	grow
8.8	*zəŋ[əɣ(i)]	*deŋeR (L)	hear
8.15	*(n)qay(i)	*kaRi (?)	say
8.21	*ɣət(s)	*peRes (I)	squeeze
8.20	*(ba)laq(i)	*belaq (?)	split
8.22	*(n)cuk(i)	*suksuk (?)	stab
8.23	*da(q)	*tudaq(?)	throw

Table 8.7 Potential PAA–PMP comparisons (13)

	<i>PAA</i>	<i>PMP</i>	<i>Gloss</i>
1.2	*buk	*qabuk	dust
1.4	*(m)put	*kabut	fog, mist
5.1	*(n)quʔan[ak]	*anak	child
5.2	*saw[a]	*qasawa	spouse
6.1	*(n)təp	*qatəp	roof/thatch
7.3	*(z)leŋ	*ma-diŋdiŋ	cold
7.7	*tu[l]	*pundul	dull/blunt
8.1	*r(a,u)wa(i)	*mañawa	breathe
8.6	*t(a,u)(m)puq	*nabuq	fall
8.7	*m[b]uk	*tu(m)buq	grow
8.20	*(ba)laq(i)	*belaq	split
8.22	*(n)cuk(i)	*suksuk	stab
8.23	*da(q)	*tudaq	throw

Table 8.8 Potential PEAA–PMP comparisons (13)

	<i>PEAA</i>	<i>PMP</i>	<i>Gloss</i>
4.3	** (m)paq	*baqbaq	mouth
5.3	**na(q)	*bana	husband
7.4	**bər, **bənər	*ma-bener	correct, true
7.5	**məz	*cemeD (?)	dirty
7.7	**tu[l]	*pundul (?)	dull/blunt
7.9	**i[ʔ]ak, **u[ʔ]aq	*ma-pia	good
7.10	**baRe(n)qət	*ma-beReqat	heavy
7.15	**[b]acaq	*ma-baseq (?)	wet
8.11	** (i)ləp	*qinep	lie down
8.14	**ntuk	*tuktuk	pound
8.18	**zaq(i), **ñjaqi	*luzaq	split
8.19	**ta[q]	*sitaq	split
8.24	** (n)kət	*hiket	tie

Table 8.9 PAN and PMP phonemic systems

<i>PAN</i>					<i>PMP</i>				
p	t		k	q	p	t		k	q
	C	c					c		
b	d	j	g		b	d	j	g	
		z					z		
m	n	ñ	ŋ		m	n	ñ	ŋ	
	N								
	S	s		h		s			h
	l					l			
	r		R			r		R	
w		y			w		y		

Table 8.10 AN-AA phonological correspondences (based on Hayes' complete set of proposed lexical correspondences)

PAN	PMP	PAA	Reconstruction #	
			PAN	PMP
pV.CV -p-	pV.CV	Ø.	8.6, 8.18	7.7, 7.9, 8.21
		-(m)p-	1.3	
	-p-	-mp-	(7.6)	
		(m)p-	8.10	3.4
		-p	4.15	
		(n)q-	(7.6)	
-p	-p	-p	8.14	6.1, 8.11
		Ø	8.3	
CVp.CVp		CVp	8.16	
tV.CV	tV.CV	Ø.	8.1, 8.17, 4.2	8.7, 8.13, 8.17, 8.23
		tV.CV	7.3, 8.7, 8.10	
		t-VC		1.3
		sV.CV	1.7	
tVC.tVC	tVC.tVC	tVC	8.8	8.14
-t-	-t-	t-	1.10, 1.11, 8.25	7.2, 8.19
		-nt-		
		-[R]-	8.17	
				1.4, 1.7, 2.5, 7.10, 7.11,
-t	-t	-t(i)	8.2	8.16, 8.24
kV.CV	kV.CV	Ø.	8.2	1.4
		kV.CV	2.2	3.4
		k-	3.4, 4.1, 8.21	
		(n)k-	8.9, 8.11	
		qV.CV	7.6	
		(n)q-		7.11, 8.15
		bV.CV		1.7
-k-	-k-	-k-		7.14, 8.3, 8.26
		-k		8.12
		(n)k-	8.1	8.24
		(n)q-	3.3	1.6, 5.4
				1.2, 3.1, 3.4, 4.1, 4.2,
-k	-k	-k	7.4	7.16, 8.7
		-ki		3.2
		-(q)		7.13
CVk.CVk	CVk.CVk	CVk	8.8	8.14, 8.22
			1.9, 1.10, 3.2,	1.2, 1.8, 5.2, 6.1, 7.2,
qV.CV	qV.CV	Ø.	(8.12), 8.14	8.11
			1.1, 1.7, (8.12),	
		qV.CV	4.10	
		[j]a.	1.2	
-q-	-q-	-q-	1.11, 4.3, 4.8	2.1, 8.10, 8.16, 7.10
		(n)q-	4.13	
		-q		7.1
			2.1, 7.5, 8.6,	1.3, 1.6, 4.4, 7.8, 7.14,

(Table 8.10 continued)

Table 8.10 Continued

PAN	PMP	PAA	Reconstruction #	
			PAN	PMP
-q	-q	-q	(8.7), 8.17, 8.18	7.15, 8.2, 8.6, 8.18, 8.20
		-(q)		7.12, 8.23
CVq.CVq	CVq.CVq	-k	(8.8)	
		CVq		4.3
CV.CV		Ø.	4.5, 5.5, (8.4),	
		tVCV	8.21	
		tVC	(8.4)	
		ntVC	6.5	
		CV.CV	8.15	
		cVCV	3.2	
-C-		-t-	4.3	
		-C-	4.6	
		-l-	3.4	
eVCV		Ø.	4.11	
		Ø.		7.5
bVCV	bVCV	Ø.	1.11, 1.12, 7.2,	2.3, 4.2, 5.3, 7.13,
			7.4	(7.15), 8.9, 8.25
		bVCV		(7.4), 7.10, (7.15),
		(bV)-	1.5, (5.2), 8.23	(7.16)
		b-	4.15	8.20
		p-	2.1	(7.4), (7.16)
		(m)p-	8.5	8.12
bVC.bVC		(m)p-		2.2
		-b-	4.13, (5.2)	4.3
		(m)b-	1.1	
		-mb-	2.1	
		m[b]-		1.1
		b-		8.7
		p-	3.5	1.2
		-mp-		1.4
		-p		8.6
		Ø.		8.27
-b		dVCV	1.4, 1.13	
dVCV	dVCV	dVCV		2.4
		(d)-		4.4
		zVCV		2.1, 7.14, 8.8
		(z)-	8.20	
		(z)VC		7.3
-d-		d-		8.17, 8.23
		(n)z-	8.19	
-nd-		t-		7.7
-d		-d	4.1 (?)	
-j-		(n)j-	6.2	2.3
		-r-	1.2	
g				

(Table 8.10 continued)

Table 8.10 Continued

PAN	PMP	PAA	Reconstruction #	
			PAN	PMP
zV.CV	zV.CV	Ø	6.3	
		jV.CV		7.8, 8.16
		ñj-		7.1
-z-	-z-	kV.CV	6.4	
		ñj-		(8.18)
		-[ñj]-	1.6	
mV.CV	mV.CV	z-		(8.18)
		Ø		3.1, 8.4
		mV.CV	4.16	4.1
-m-	-m-	m-	3.2	7.5, 8.2
		-m-	8.26, 4.9, 5.1,	
		-(m)p-	6.1	
-m	ñV.CV	-m	1.7	
		rC.CV	1.13, 6.3, 8.13	7.2, 8.5, 8.13
		nV.CV		8.1
-n-	-n-	tV.CV		7.6
		n-	5.4	5.3
		-n-		5.1, 7.4
-n	-n	-n		8.9
		l-	1.4, 1.9	3.1, 8.11
		-l	1.1	2.4, 8.17
ŋV.CV	-ŋ-	-m[u]s	6.2	
		Ø	6.1	
		-ŋ-		8.8
-ŋ-	-ŋ	-ŋ	4.5	2.2, 7.6
		-ŋ	4.3	7.3, 8.25
		lV.CV	8.24	
-N-	-N-	-l	5.3	
		-n-	8.4	
		(c,s)[r]-	(8.14)	
-N	sV.CV	(n)[r]-	(8.14)	
		-l	4.2	
		Ø	3.3, 8.3, 8.10	
sV.CV	sV.CV	sV.CV	3.6	
		[s]V.CV	(1.3)	
		sVC	4.7	
-S-	-S	[s]VC	(1.3)	
		cV.CV	4.12	
		gV.CV	7.5	
-S	sV.CV	-h-	2.2	
		q-	8.17	
		-s(i)	7.3	
sV.CV	sV.CV	Ø		8.19
		sV.CV	1.8	5.2
		[n]sV		

(Table 8.10 continued)

Table 8.10 Continued

PAN	PMP	PAA	Reconstruction #	
			PAN	PMP
		cV.CV		8.3
sVC.sVC		sVC	(4.22)	
	sVC.sVC	cVC	(4.22)	8.22
	-s-	s-		7.13
-s-		c-	3.1	1.10
		-c-	4.12	7.15
-s		-c(i)	1.7	1.5
	hV.CV	Ø.		8.24, 8.27
-h-		(n)s-		4.2
	-h-	-[q]-		
		-[x]-	5.2	
-h		-s	8.4	
		-[q]	8.26	
IV.CV	IV.CV	Ø.	3.5, 4.1, 4.13	2.3, (3.3), 8.18
		!CV.CV	4.9, 8.6	
			1.5, 3.2, 3.6,	
-l-	-l-	-l-	4.10, 6.4	1.7, (3.3)
		-[l]-	8.5	1.6, 8.25
		!l	6.5	
		!l	4.5	1.5, 4.4, 8.20
		[]l-	8.12	
-l		-l	1.10	
		-[l]	7.6, 8.9	7.7
	rV.VC	Ø.		(7.16)
		-r-		(7.16)
	CVr.CVr	CVr		1.9
	-r	-r		7.4
RV.CV	RV.CV	yV.CV	7.1	1.1
		yVC	6.1	7.6
-R-	-R-	-y-	8.7	2.5
		-[y]-	1.8	
		y-	7.4, 8.2	7.12, 8.21
		-y	8.20	8.4, 8.15
		-R-	4.14	7.10
		R-	6.3	
		-r[]	8.23	
-R	-R	-y	4.7, 8.12, 8.19	8.8
		-R	3.2, 3.6	
		-[R]	4.13	
		-[R]-	7.2	
		-y		2.6

Table 8.11 Results of the evaluation of Hayes' basic vocabulary comparisons

	<i>Probable</i>	<i>Possible</i>	<i>Weak</i>	<i>Rejected</i>
A. PAA-PAN	9	9	2	12
B. PAA-PMP	3	5	9	8
C. PEAA-PAN	3	5	0	5
D. PEAA-PMP	2	3	1	7
Totals	17	22	12	24

The Appendix to this chapter provides my evaluation of what I consider to be the potentially corresponding forms among the Hayes' list of basic vocabulary. Table 8.11 summarises the results.

The morphosyntactic evidence for Austric

Ross (2000: 447), in his careful review of my most recent paper on the subject (Reid 1999), wondered whether I was no longer satisfied that the morphosyntactic evidence I had cited in my earlier paper was still viable as evidence. In that I was presenting 'new evidence for the hypothesis' I did not think it was necessary to restate the old evidence.¹ However, for those who may not be familiar with the earlier work in this area, I will restate it here.

Evidence given in my 1994 paper includes:

- 1 The AA causatives *pa-/<ap> and *ka- are considered to correspond to AN causatives *pa-, *ka- and *paka-.
- 2 The AA agentives *<um> and *ma-/<am> are considered to correspond to AN agentives *mu-/<um> and *maRa-.
- 3 The AA instrumentals *<an>, <in> are considered to correspond to EF instrumental *paN-, and AN nominalising affix *ni-/<in>, respectively.
- 4 The AA objective *-a is considered to correspond to AN objective *-a.
- 5 Evidence from Sora, Khasi, Nancowry and Car Nicobarese suggest a PAA attributive linker *(n)a corresponding to the AN 'ligature'*(n)a
- 6 The Nicobarese determiners marking case of NPs *ʔin*, *ʔan*, *nun*, etc., appear to have developed by the same well-known grammatical processes that have brought about the nasal final determiners in many AN languages, that is, by the fusion of a reduced form of the ligature *na. What is important here is that the Malayic languages reflect a PMP (my use of the term, not Blust's) innovation *na > *ŋa, hence Tag. *ang*, *nang* Kawi *ang*, Malay *yang*, etc., so that the Nicobarese forms could not have been borrowed from sailors speaking a Malayic language.
- 7 Evidence from Nicobarese, Old Khmer, Khmu and Mal suggest a PAA *ta 'locative' preposition corresponding to PAN 'locative preposition, demonstrative' *ta.
- 8 Car Nicobarese *ʔi* 'locative preposition' corresponds to PAN *i 'locative preposition'. Note that although reflexes of this preposition are found all the way to Proto-Polynesian, the Proto-Malayic locative preposition is

reconstructed as *di (Adelaar 1992), and is a reflex of PAN *di ‘locative preposition, demonstrative’, so if borrowed, the Nicobarese locative preposition could not have come from sailors speaking a Malayic language.

Ross (2000: 446–7) neatly summarises the morphosyntactic evidence that I had presented in my 8ICAL paper as follows:

- 9 The Nicobarese causative verbal infix <um> is taken to be cognate with PAN infix *<um> that marked a verb as an unergative intransitive and formed deverbal nouns expressing non-agentive causers (e.g. Bontok *s<um>akit* ‘that which makes [someone] sick’). (Refer also Schmidt 1916.)
- 10 PAN and PAA are both taken to have been ergative with a contrast between nominative and genitive pronouns, the genitive denoting both possessor and transitive agent.
- 11 The PAN and PAA first-person singular pronouns appear to be cognate.
- 12 Nominative pronouns in both PAN and PAA are taken to have been prefixed with *a-.
- 13 Ruc, a conservative language of the Vietic branch of AA, has a dative prefix *pa-* that appears to be cognate with PAN *pa ‘go’.
- 14 A non-proximal demonstrative *en is reflected in both AA and AN languages.

Several arguments have been raised in recent years in an attempt to find alternatives to the morphological comparisons cited earlier.

Borrowing is the primary explanation that has been proposed. Sagart (email comm. 2001) states, ‘I have come to the conclusion that the “accidental/involuntary action” prefix *ta-* in AA languages of Vietnam: Pacoh, Chrau, Katu and Bahnar, is borrowed from Chamic. So this argues that transmission of morphology from AN to AA is possible’.

There is little question that morphological processes can be borrowed between languages of different families. Whether or not it is possible for ‘morphemes’, that is, meaningful phonological units which constitute part of a word, to be borrowed without their host words also being borrowed is a matter still open for discussion. But whatever the answer to that question, the morpheme *ta-* itself has not been proposed as evidence for Austric, and what is more important, the morphemes that have been proposed have such an extensive distribution, within the AA family, that no reasonable explanation can be given for either the time of borrowing or the possible source language.

The possibility of borrowing seems likely in the case of the strong Nicobarese morphosyntactic similarities with AN, where it is assumed early AN sailors may have made frequent landfall, perhaps in some cases staying, intermarrying and influencing the local language. But there remain two strong barriers to the acceptance of this position. One is that several of the proposed comparisons between Nicobarese languages and AN are not limited to Nicobarese, but are found across wide areas of the AA family. Comparisons in some cases (especially <um> and <in>, however, are clearest with Nicobarese because other EAA languages have either lost the form (in the

case of verbal suffixes) or modified them as the result of the strong areal influence of Chinese. Moreover it is clear that if Nicobarese borrowed the suspect forms, it could not have been from a language of the Malayic family, because by Proto-Malayic times the forms had been either lost or changed from the earlier forms that I claim are reflected in Nicobarese (as noted earlier (6) and (8)).

The claim has also been made (Sagart email comm. 2001) that the vowels of the **mu*-/<um> and **ni*-/<in> affixes

are more or less colored by the main consonant, they could be secondary. The vowel is essentially an epenthetic schwa which serves to break the consonant cluster formed by prefixed *m*- or *n*- and the root initial, or the root initial plus infixes *-m*- and *-n*-. In both cases the prefix or infix tends to color the schwa: you get rounding with *m*, and a high-and-some-what fronted vowel with *n*. It is possible that in each particular word the root initial also plays a role coloring schwa.

Whether or not this may ultimately be the source of the vowels in these affixes, the evidence that we have from Nicobarese, and from AN languages suggests that in these languages at least, the vowels of the affixes were full vowels, and not epenthetic. PAN alone has a number of reconstructions which show *m* followed and/or preceded by vowels other than *u* (e.g. **mimah* ‘drink’, **SimaR* ‘fat, oil’, **maCa* ‘eye’, **gemgem* ‘hold (in fist)’, etc.), and *n* followed and/or preceded by vowels other than *i* (e.g. **tanek* ‘cook’, **tenem* ‘sea, saltwater’, etc.). Similar data could be drawn from Nicobarese as well as from Munda languages, and these are the only places that one could look in AA for evidence to support the thesis.

With respect to my claim that the AA instrumentals *<an>, <in> are considered to correspond to the EF instrumental **paN*-, and the AN nominalising affix **ni*-/<in>, respectively, the argument has been made that the functions of the AA and AN affixes are not close enough, the former being ‘instrumental’ or ‘agentive’ (Thurgood 1999) while the latter was an ‘objective’ nominalisation specifying the result of the action of the verb. However, AA data not only shows ‘instrumental’ nominalisations, but nominalisations of the AN type. Thurgood (1999: 245) quotes Banker as claiming that the Bahnar infix <*ən*> sometimes means ‘the result of a verbal action’, so that *bât* ‘to make a dam’, becomes *b<ən>ât* ‘a dam’. It is probable that the agentive function of the affix was also present in PAN, given the form **C<in>aqi* ‘guts’, which can only be interpreted as meaning ‘that which produces **Caqi* “faeces”’. That the same affix can be reconstructed with these two apparently quite different functions should not be surprising, given the fact that forms such as *b<ən>ât* ‘a dam’ are potentially ambiguous between ‘the object that is the result of damming’, and ‘the object that dams’.

One of the facts which is of high value in supporting the morphological comparisons between AA and AN languages, is not simply the forms and their functions, but also the apparently unique phonological process (not, as far as I know, reported anywhere else in the world), whereby the consonant of the affix and the initial consonant

of the root metathesise, producing alternation historically not only between the *mu-/<um> and *ni-/<in> affixes, but also between the *ma-/ <am> and *pa-/<ap> affixes, with varying distributions in both families, and clearly with the original metathesis (producing infixes) reversible, so that in some daughter languages the prefixal forms reappear either alone, or in alternation with the infix, depending (usually) upon the manner of articulation of the root-initial consonant.

Ross (2000) has questioned the identification of *mu-/<um> as a causative nominaliser in PAN, suggesting that it is by no means certain that it was a nominaliser at all, thereby questioning its functional association with the corresponding PAA forms. If PAN was anything like most present-day AN languages, whether or not it functioned primarily as a nominaliser or not, verbs that were formed with it, would also have been zero-derived as nouns, when appearing as the heads of noun phrases, just as the Bontok form *s<um>akit* can mean either ‘to make one sick’ when occurring as a verb, or ‘the thing that makes one sick’, when occurring as a noun.

He further questions my identification of *a- as a nominative marker in PAA and PAN with the observation that in Taoih, this prefix is found not only on subject pronouns, but on dative pronouns as well. Of course, if PAA was ergative, as I claim, what is today a dative pronoun in an accusative language (as Taoih probably is) would have been the grammatical subject of a transitive sentence at an earlier stage and would have been marked as nominative, as it is in AN languages that are ergative, for example, Bontok,

<i>Agtam</i>	<i>sak-en</i>	<i>si</i>	<i>itab</i>
give.Gen.2s	Nom.1s	of	beans
Give me some beans.			

To me this is further evidence in support of the ergative nature of PAA. Compare also the *?an* subject marking of some NPs in Nicobarese. Ross questions my claim that the initial *a- of the PAN nominative pronouns *aku, *aken, *aten and *amen was the original nominative marking component of these forms, by claiming that genitive forms such as Seediq *n-aku*, Pazeh *n-aki*, Thao *n-ak*, *n-am*, Amis *n-ako*, etc., showed that even in PAN the *a- was part of the pronominal root and could occur as a genitive. The hyphens in these forms represent Ross’ analysis of the forms, as having an initial *n-* genitive marker. I claim however that such forms are better analysed as: Seediq *na-ku*, Pazeh *na-ki*, Thao *na-k*, *na-m*, Amis *na-ko*, etc., an analysis which is more consistent with general patterns of genitive noun formation in a wide range of AN languages, including Talubin Bontok in which *nak* ‘my’, *nam* ‘your’ have independently developed from a combination of *na* ‘the, non-referential noun’ plus genitive pronominal endings.

Conclusion

Ross claims that in order to be convinced of the validity of Austric, he would need either a substantial quantity of regularly corresponding cognates, or a seemingly cognate paradigm of grammatical morphemes. It would be great if we were able

to provide either one or the other, but at the time-depths we are looking at, and the imperfect state of our knowledge of AA languages, and the extremely limited amount of reconstruction that has been done in the family (compared at least to the AN side), what I have presented in this chapter is at present the best we can do. I believe that the number of apparent cognates cited here between PAA (and PEAA) and PAN (and PMP) from the area of basic vocabulary, come close to providing such a convincing body. The hope of providing seeming cognate paradigms of grammatical morphemes comes closest with the sets of what were probably originally demonstrative nouns, but which in both families have grammaticised into a wide range of determiners, ligatures, prepositions and the like. There is no question that the range of forms is there, including *a, *ta and *na, with corresponding functions on both sides of the family. It is unlikely, however, that it will be possible to find any paradigm of verbal morphology (which I suspect is what Ross is looking for), because I don't think there was much of this in early AN. The paradigm of so-called 'focus' morphology, even if it was present in PAN (which I think is doubtful), is clearly a catch-bag of prefixes, infixes and suffixes which must have existed in pre-AN times as probably nominalising affixes, but never in any sense constituting a paradigm.

The evidence then is not as convincing as one would like, but as Diffloth said in 1994, the evidence 'is undoubtedly there', and I believe it is considerably stronger now, than it was then. The evidence is for a genetic relationship, but is it evident that the families in question descended from a common immediate ancestor, Proto-Austrie? With the accumulation of evidence presented by Sagart in this volume and elsewhere, that AN can also be shown to be genetically related to the Sino-Tibetan family of languages, and his claim (Sagart p.c.) that some of the lexical items and affixes claimed to be shared by AN and AA are found also in Sino-Tibetan languages, the possibility exists that the relationship between AA and AN is more remote than earlier considered. The concept of 'Austrie' as a language family may eventually need to be abandoned in favour of a wider language family which can be shown to include both AN and AA language families, but not necessarily as sisters of a common ancestor.

Appendix

Evaluation of potentially corresponding forms among the Hayes list of basic vocabulary

A. PAA–PAN

1. Probable

ashes

PAA *qabuh

PAN *qabu

Pacoh *abóh*, Chrau *vuh* 'ashes',

Bonda *bu?* 'to smoke'

Comments: Accepted by Diffloth as possible (1994: 313). Restricted to Katuic (Pacoh) and Bahnaric (Stieng and Sre), and a Munda cognate with questionable semantics. Probably not borrowed.

dog

PAA *cu(q)

PAN *asu

Bonda *gusɔʔ*, PW *sɔʔ, VN *chó* ‘dog’*Comments:* Accepted by Diffloth as probable (1994: 313).**snake**

PAA *[su](l)aR

PAN *SulaR

Kharia *lur*, Sora *loʔor*, (CF *lor*) ‘a kind of snake’, Bahnar ‘*bih tep-lar*’ ‘a very small snake that is extremely poisonous’*Comments:* Possible final syllable reflexes in Munda and Bahnar.**belly**

PAA *taʔal, *tiʔal

PAN *tiaN

Sora *taʔal* ‘spleen’, Thavung *khaʔal* ‘belly, stomach, abdomen’, Pacoh *acheal* ‘heart’*Comments:* Final syllables match from Munda through the Muong and Katuic comparisons. Semantics are acceptable, and PAA *l corresponds to PAN *N in several cases.**eye**

PAA *mə(n)ta(q)

PAN *maCa

Kharia (V250) *mɔʔd*, PVM *mat ‘eye’, Proto-Plang *hak^l-kitaʔ^l ‘eyebrow’*Comments:* Accepted by Diffloth as probable (1994: 317). Note Proto-Plang (Waic) *hak^l-kitaʔ^l ‘eyebrow’ Lit. hair-eye, which supports the final PAA syllable.**father**

PAA *(qa)ma(ma)

PAN *t-ama

Santali *mama* ‘maternal uncle’, Katu *ama*, Pacoh *a-ám* ‘father’, Bahnar *ma* ‘younger brother of father or mother’*Comments:* Widely distributed with appropriate phonology and semantics, but suspect as a possible nursery word.**mother**

PAA *(na)na

PAN *t-ina

Kharia *nana* ‘elder sister’, Bonda *tuna* ‘younger sister (addressed by a brother), wife’s younger brother’s wife’, Sedang *na* ‘older sister, cousin’*Comments:* Widely distributed with appropriate phonology and semantics, but suspect as a possible nursery word.**rotten**

PAA *yok

PAN *ma-buRuk

Kharia *lorog* ‘to rot, decay’, VN *ruc* ‘be rotten’, NK *phròok* ‘spoiled’*Comments:* Phonologically plausible with possible reflexes in Munda, Monic and Viet–Muong groups.**buy**

PAA *pə[ɿ]i

PAN *beli

Khasi *pli* ‘change’, MUK *pál*, *pánh* ‘sell’, Kharia *pətay* ‘fix price, bargain’
Comments: The Khasi, Muong and Katuic forms appear to be cognate, and probably correspond to the PAN form.**2. Possible****salt**

PAA *qa(m)puc(i)

PAN *timus

Pareng *b.sut*, Kuy *pos*, Jehai *mpɔj* ‘salt’*Comments:* Possible cognates in Munda, Pearic, and Aslian.**smoke**

PAA *(m)pəʃ

PAN *qebel

Sora *poro* ‘become smoky’, *mor* ‘spread as smoke’, Stieng *pôr* ‘smoke out of a hollow tree’

Comments: Possible cognates in Munda and Bahnaric.

egg

PAA *te|oR

PAN *qiCeluR

Juang *susuter*(ɔ), Pacoh *tireal*, *tirōl*, PVM *t(ə)lur? ‘egg’

Comments: The Munda comparison is doubtful. The Katuic and Viet–Muong comparisons seem more secure.

male/man

PAA *(n)qalay

PAN *ma-RuqaNay

Bonda *laibu?* ‘male pig’, Pacoh *alay*, Stieng *clay* ‘brother-in-law’

Comments: The Munda, Katuic and Bahnaric forms possibly correspond. If they do, they probably correspond to the PAN form.

rat

PAA *b[əw]

PAN *labaw

Bonda *gubu* ‘a kind of rat’, Riang (Black) *kəbu*¹ ‘rat, mouse’, Mah Meri (Bes. K.L., R33) *kanē* ‘*rəbu* ‘mouse’

Comments: The Munda, Bahnaric and Aslian forms seem to be cognate. The final syllable is a possible comparison with the PAN form. None of the AA initial syllables correspond with the PAN initial syllable.

head

PAA *(n)qo|u(q)

PAN *qulu

Bahnar (PB) *kōl*, Jeh *kāl*, Mal *kliq* ‘head’

Comments: Reasonable phonological correspondences between Bahnaric

and Aslian, possibly corresponding with the PAN form.

shoulder

PAA *(ba)Ra(q)

PAN *qabaRa

Kharia *taran*, Theng *blah* ‘shoulder’, Khasi *ta-bla* ‘shoulder piece of animal’

Comments: The Khmuic and Khasi comparisons appear good, and probably correspond with the PAN form, but the Munda term is questionable.

hold (in fist)

PAA *(n)kəm

PAN *gemgem

Kensiu *cəkam*, VN (*gəm>) *câm* ‘hold’, Sora *kum-si*: ‘hold in one’s fist, hold a handful’

Comments: The Aslian, Vietic and Sora are possible cognates, and if so, probably correspond well with the PAN form.

yawn

PAA *(can)qap

PAN *ma-Suab

Santali (V68) *aygɔ*^b, PM *sɲʔaap, VN *ngáp* ‘yawn’

Comments: The Munda, Vietic and Mon forms are possible cognates. They possibly correspond to the PAN form.

3. Weak

meat/flesh

PAA *c(i, ə)ci

PAN *Sesi/isi

Sora *sissid*, VN (*ñsic>) *thit* ‘flesh, meat’, PM *sac ‘fruit, nut, berry, acorn, pod’

Comments: Insecure phonological correspondences.

swim

PAA *la(n)[ʔ]oy

PAN *Naŋuy

Mundari (K519) *oiyar*, Ruc, *loy*, Riang Lang *_ɲɔy* ‘swim’
Comments: The Munda form is probably not cognate with the eastern forms. The Palaungic and Ruc forms are likewise doubtful comparisons.

4. Rejected

breast

PAA *n[s]uq

PAN *susu

Bonda *daʔtu kuĩ*, PW *tis ‘breast’, Semai *ntoh* ‘chest’

Comments: Phonological correspondences don’t work.

neck

PAA *(n)qe[R]

PAN *liqeR

MUK *kel* ‘neck’, Pacoh *cõl* ‘wear around neck’, Sengoi *kelkeil* ‘ankle, wrist’

Comments: The Muong form possibly corresponds with the PAN form, but the Katuic and Aslian forms are semantically doubtful.

female/woman

PAA *a[x]i, *bu[x]i, *mpa[x]i,

PAN *bahi

Pacoh *a-i* ‘mother’, Kharia (K349) *bui* ‘girl’, Mon *imbay* ‘elder brother’s wife, husband’s elder sister’

Comments: Unconvincing phonological and semantic correspondence.

big

PAA *(can)ɣaya(q)

PAN *ma-Raya

PW *ra ‘big’, Theng *yaʔ* ‘far’, Sora (V40) *saya: j-ən* ‘be at a distance’

Comments: The Proto-Wa form is possibly cognate, but the semantics and

phonology of the Khmuic and Munda forms are unconvincing.

shy/ashamed

PAA *g(i,a)haq(i)

PAN *ma-Siaq

Bonda *gĩak* ‘shame’, PW *[gac] ‘ashamed, shy’, Stieng *haas* ‘feel ashamed, bashful’

Comments: Phonological correspondence with the PAN form is unlikely.

bite

PAA *(n)ɣat(i)

PAN *kaRat

Sora (V334) *gad* ‘cut’, *raj* ‘cut into small pieces as wood’, PM *rac ‘cut with a sickle, reap’, Katu *karóoch* ‘cut kernels off’

Comments: Phonology may be possible, but semantics unlikely.

dream

PAA *(m)pe(qi)

PAN *Sepi

Chrau *vĩq*, Katu *bách* ‘lie down, sleep’, Sora *mimid̥* ‘sleepy’

Comments: Phonological correspondences don’t work.

hunt

PAA *(n)[r]op, *(c,s)[r]op

PAN *qaNup

Jeh *rũp* ‘catch, seize’, Khasi *kynrup* ‘pounce upon, seize’, Bonda *sɔp* ‘hold, catch’, Stieng *choop* ‘hunt’

Comments: Only the rhyme works.

sleep

PAA *(n)zoɣ

PAN *tuduR

Pacoh *chur* ‘sleepy or sad eyes’, Birhor (V111) *durum* ‘to sleep’, Khmu’ *hmdir* ‘to snore’

Comments: Unlikely semantics and phonology.

steal

PAA *ka|aw *kum|aw

PAN *Cakaw

Nicobar *kə̀lɔː-hə̀ŋə̀* ‘steal’, Mundari (V242) *kumɾu*, Santali *kombɾo* ‘thief, theft, steal’

Comments: Phonological correspondences unlikely.

suck

PAA *s[e]p, *(ñ)c[e]p

PAN *sepsep

VN *top* ‘sip’, Mundari (V354) *si’b* ‘to smoke’, Pacoh *dyép* ‘suck’

Comments: Phonological correspondences unlikely.

vomit

PAA *taq

PAN *utaq

PM *taaʔ ‘vomit’, Bonda *taʔmi* ‘sneeze’, Semelai *tahtəh* ‘to spit’

Comments: Unlikely semantics.

B. PEAA–PAN*1. Probable***lake**

PEAA {**|əw} (Hayes)

PAN *danaw

Chrau *tanlô* ‘lake, pond’

Comments: Hayes notes that the lateral in the Chrau reflex he cites (*tanlô* ‘lake’) suggests that the form could not be a borrowing from Chamic.

head louse

PEAA **k[o]t(i)

PAN *kuCu

Katu *kóót* ‘lice, fleas, bugs’, Khmer *sarikæc* ‘bedbug, flea’, Chrau *sicäch* ‘tick’

Comments: With apparent cognates in Katuic, Bahnaric and Khmer, this is an attractive set to correspond with the PAN form.

afraid

PEAA ***(n)k[o]t

PAN *ma-takut

Jeh *kokiat* ‘detest, hate’, Khmer *kot* ‘hold in awe’, Mon *takuit* ‘take fright’

Comments: With apparent cognates in Aslian, and both Mon and Khmer, the phonology looks reasonable and the semantics plausible.

*2. Possible***wood**

PEAA **k(i,a)hi(uq)

PAN *kaSiw

PW *khoʔ ‘tree’, PM *chuuʔ ‘tree, wood’, Semai *jəhuʔ* ‘tree’

Comments: An interesting set of forms which look as though they may well be cognate but the phonological correspondences are uncertain.

fruit

PEAA ***(m)b[oʔa]q

PAN *buaq

Kensiu *kəbəʔ*, Sabum *kəmʔ* ‘fruit’, Stieng *moq* ‘type of small fruit’

Comments: The Aslian and Bahnaric appear to be cognate, and there is the possibility that they correspond also with the PAN form.

new

PEAA **qi[R]u(q)

PAN *ma-baqeRu

PW *croʔ ‘new’, Bahnar *chrêu* ‘strange’

Comments: The Waic and Bahnaric forms seem to be cognate. At least their final syllable may correspond to the PAN form.

burn

PEAA **[ʔ]us(i), **t[u]nus

PAN *CuNuh

Katu *pa-óh* ‘cook’, Khmer *’us* ‘firewood’, Bahnar *tonuh* ‘hearth’

Comments: The Bahnar form is suspiciously similar to the PAN form. But the form is not reconstructed for Proto-Chamic and is probably not a borrowing. Other AA languages show a lateral corresponding to PAN *N.

cut (wood)

PEAA **tayaq, **tayak

PAN *taRaŋ

Rengao *chrā* ‘split, divide, crack open’, Sre *trac* ‘shave’, Pacoh *trəŋ* ‘chop’

Comments: The Bahnaric forms may well be cognate with Pacoh. They may correspond to the PAN form.

3. Weak

rain

PEAA **ka[ŋj]al

PAN *quzaN

Brou *cuyal*, Old Mon *kyāl*, Khmer *khya’l* ‘wind’

Comments: Although the AA forms are probably cognate, only the rhyme appears to correspond with the PAN form.

stone

PEAA **tamuq (Hayes); PMK *təm(o)ʔ (Diffloth)

PAN *batu

Khasi *máw, PW *smoʔ, PM *tmɔʔ ‘stone’

Comments: Diffloth (318) notes this as a possible correspondence. The evidence is weak.

burn

PEAA **[ʔ]us(i), **t[u]nus

PAN *CuNuh

Katu *pa-oh* ‘cook’, Khmer *’us* ‘firewood’, Bahnar *tonuh* ‘hearth’

Comments: The Bahnar form is suspiciously similar to the PAN form. But the form is not reconstructed for Proto-Chamic and is probably not a

borrowing. Other AA languages show a lateral corresponding to PAN *N.

cut (wood)

PEAA **tak, **tek

PAN *tektek

Katu *ntaak* ‘chop’, Pacoh *tích* ‘chop firewood’, Rengao *kotək* ‘snap, break, cut skin’

Comments: The consonants appear to correspond, but the forms are probably onomatopoeitic.

4. Rejected

moon

PEAA **b(i,a)l

PAN *bulaN

Katu *baraal* ‘pale’, Bateg Deg *bəyəl* (*r > y) ‘white’, Bahnar *monhal* ‘very bright light or sunshine’

Comments: Unlikely semantics.

sand

PEAA **lay

PAN *qenay

VN *lây* ‘miry, swampy, marshy’, Bahnar *lai* ‘mound of dirt’, Nyah Kur *lɛɛ, rɛɛ* ‘ore, mineral’

Comments: The may phonology be possible, but the semantics are unlikely (none means ‘sand’).

back

PEAA **ko[d(i)]

PAN *likud

Boriwen *kiat* ‘back’, Halang *kuyq* ‘small of the back of the head’, Jeh *kung kuyq* ‘back of head’

Comments: The Bahnaric are probably not cognate with Boriwen, nor with the PAN form.

liver

PEAA **p[a]le(q)

PAN *qaCay

Mon *pli* ‘spleen’, PVM **ple?*,
Sengoi *pele* ‘fruit’

Comments: The AA forms may be cognate, but they certainly don’t seem to relate to the PAN form, either in phonology or semantics.

old (people)

PEAA ***ti(n)qas(i)*

PAN **ma-tuqas*

Khmer *cā’s*, Pearic *čhu:s* ‘old’, Katu *takóh* ‘grown’

Comments: The phonological correspondences between AA and the PAN form don’t work.

choose

PEAA ***(u)laqi*

PAN **piliq*

PW **ras*, Pacoh *rôih*, Semai (Serau, C120A) *chenlas* ‘choose’

Comments: Impossible phonology.

see

PEAA (**)(*s*)[*R*]ai

PAN **kita*

Pacoh *lây*, Chrau *sây*, VN *thây* ‘see’

Comments: Impossible phonology.

sew

PEAA ***(n)qiq*

PAN **taSiq*

Pacoh *ěh*, Katu *jih*, Sengoi

ceik ‘sew’

Comments: Impossible phonology.

C. PAA-PMP

1. Probable

dust

PAA **buk*

PMP **qabuk*

Bonda *tubok’/tubuk’* ‘earth’, Chrau *vôq* ‘mud’, Mon *khabuik* ‘fine powder or dust’

Comments: Reasonable semantics, and good (final syllable) correspondences from Munda through Bahnaric and Mon.

roof/thatch

PAA *(*n*)təp

PMP **qatep*

Khasi *tap* ‘to cover’, Mundari (V3)

da’b ‘cover a roof, thatch’, Palaung

dļp ‘to cover, thatch’

Comments: Good phonological and semantic correspondence across AA, with regular correspondences to the PMP form.

split

PAA *(*ba*)laqi(i)

PMP **belaq*

Katu *blah* ‘split’, Kharia (V304)

la’j ‘slice’, Khmer *-la’s* ‘separate, detach’

Comments: The Katuic form is cognate with reconstructed forms in PMong **blah*, as well as in three branches of Bahnaric (Thurgood 1999: 284) so is probably not a Chamic borrowing.

2. Possible

fog/mist

PAA *(*m*)put

PMP **kabut*

Sora (V384) *umod-ən* ‘fog, mist’,

Khmu’ (*hm*)*puut* ‘clouds, fog’

Comments: Restricted distribution to Munda and Khmuic, but semantics are reasonable, and phonological correspondences possible.

spouse

PAA **saw[a]*

PMP **qasawa*

Kharia (K535) *sou* ‘husband’, Katu

sasaau ‘father’s cousins, sister’s husband, father’s sister’s children’, Proto-

Semai **bnsaaw* ‘wife’s elder brother’

Comments: The phonological comparison is attractive, and the semantics possible.

fall

PAA *(a,u)(m)puq

PMP *nabuq (?)

Mundari (K149b) *tombɔʔ* ‘fall forwards’, Katu *tampoh* ‘drop’, Bahnar *puh* ‘slip, fall into a hole’

Comments: The AA forms appear to be cognate, and possibly relate to the PMP form.

stab

PAA *(n)cuk(i)

PMP *suksuk

Sora *suj*, VN *chọc* ‘pierce’, Sengoi *cok* ‘stab, pierce’

Comments: Possibly cognate.

throw

PAA *da(q)

PMP *tudaq

Santali (V173) *lebda* ‘throw’, Khasi *pda* ‘throw to the farthest distance possible’, Chrau *randāh* ‘throw down’

Comments: The final syllable of the AA forms could correspond to the PMP form.

4. *Rejected*

child

PAA *(n)quʔan[ak]

PMP *anak

Santali (V205) *hɔn* ‘son, child’, PM *kɔɔn ‘child, offspring, young (animals)’, Mintil *ʔawaʔ* ‘child’

Comments: The Aslian form would be good were it not for the medial consonant that doesn’t correspond. The other forms don’t correspond at all to the PMP form.

cold

PAA *(z)leʔ

PMP *ma-diŋdiŋ

Kharia (K208) *rayga*, VN

lanh ‘cold’, Khmer *srei* ‘to cool’

Comments: Phonological correspondence lacking.

breathe

PAA *r(a, u)wa(i)

PMP *mañawa

Mundari (K537) *rowa*, Sengoi *ruai* ‘soul, spirit’, Pochon *rvai* ‘soul’

Comments: Phonological correspondence uncertain.

grow

PAA *m[b]uk

PMP *tu(m)buk

Kharia (V286) *muʔ* ‘come out’, Muk

mɔc ‘grow, come up’, OM

mok ‘appear’

Comments: The Kharia and Old Mon forms appear to be cognate, but the semantics of the Muk form is only questionably related. Insufficient evidence to establish a correspondence with the PMP form.

split

PAA *(ba)laq(i)

PMP *belaq

Katu *blah* ‘split’, Kharia (V304)

laʔ ‘slice’, Khmer *-laʔ* ‘separate, detach’

Comments: The Katuic form is cognate with reconstructed forms in PMong **blah*, as well as in three branches of Bahnaric (Thurgood 1999: 284) so is probably not a Chamic borrowing.

D. PEAA-PMP

1. *Probable*

mouth

PEAA ***(m)paq

PMP *baqbaq

Pacoh *piaq* ‘mouth, opening, end of river’, Mah Meri *pak*, Sengoi *mpak* ‘mouth’

Comments: Although corresponding forms are limited to Aslian and Pacoh, it is unlikely because of their shape that they were borrowed from either Malay (in the case of Aslian) or Chamic (in the case of Pacoh). Note PChamic *babah ‘mouth’ (Thurgood 1999: 283). Probably cognate with the PMP form.

dull/blunt

PEAA **tu[ʔ]

PMP *pundul

Pacoh *túl múl* ‘(expressive) of blunt end’, Bahnar *túl* ‘dull, not pointed’, NK *thuul* ‘blunt, not pointed’

Comments: Possibly cognate forms with distribution in Katuic, Bahnaric and Monic branches, appear to correspond well with the PMP form.

2. Possible

husband

PEAA **na(q)

PMP *bana

Thavung *nAA*² ‘mother’s younger brothers’, Sengoi *menah* ‘parent’s younger brother’, Bahnar *nă* ‘parent’s elder sibling’

Comments: Possible phonology, semantics questionable.

pound

PEAA **ntuk

PMP *tuktuk

Khmer *tuk* ‘beat, pound’, PM *kndok ‘pound (earth)’, MUK (*duk<) *tuc* ‘to chisel’

Comments: Plausible phonological and semantic correspondence but possibly onomatopoeic.

tie

PEAA ***(n)kət

PMP *hiket

Cua *takoot* ‘tie a knot’, VN *côt* ‘tie up, chain’, Pearic *khɔ:t* ‘tie’

Comments: The AA forms are probably cognate, and possibly correspond to the PMP form. The vowel however may be problematic.

3. Weak

split

PEAA **ta[q]

PMP *sitaq

Chrau *tăh* ‘slit open, cut up’, Stieng *tah* ‘disembowel’

Comments: The AA forms clearly cognate, but semantics don’t match well with the PMP form.

4. Rejected

correct, true

PEAA **bər, **bənər

PMP *ma-bener

Sengoi *bor* ‘good, fine, beautiful’, Bahnar *bernor* ‘goodness, righteousness, true’, Pacoh *nnór* ‘happy’

Comments: Too restricted distribution in AA, questionable semantics.

dirty

PEAA **məz

PMP *cemeD (?)

Mon *mih* ‘body dirt’, Mintil *kamah* ‘dirty’

Comments: Too restricted distribution in AA, questionable phonology.

good

PEAA **i[ʔ]ak, **u[ʔ]aq

PMP *ma-pia

VN *uóc* ‘to desire, wish for, hope for’, Khasi *kwah* [kawʔ] ‘wish for’, Jeh *wă* [waʔ], ‘want, like, be fond of, desire’

Comments: Insufficient phonological correspondence.

heavy

PEAA ****baRe(n)qət**

PMP ***ma-beReqat**

Cua *parêq*, Chrau *gât*, Mendriq (Pang. Gal., H68) *hênjut* ‘heavy’

Comments: Insufficient phonological correspondence.

wet

PEAA ****[b]acaq**

PMP ***ma-baseq**

Pearic *pəča* ‘k’, Chrau *suh*, Tampuan *tšətšuih* ‘wet’

Comments: Insufficient phonological correspondence.

lie down

PEAA *****(i)ləp***

PMP ***qinep**

Nha Heun *plíp*, Thavung *kñiip*, Bahnar *’nhíp* ‘close eyes’

Comments: Semantics not close enough.

spit

PEAA ****zaq(i), **ñjaqi**

PMP ***luzaq**

Bahnar *kosoh*, OM *ksas* ‘spit’,

Khmer *khjā* ‘k’ ‘spit out’

Comments: Insufficient phonological correspondence.

Abbreviations

AA	Austro-Asiatic
AN	Austronesian
EAA	Eastern Austro-Asiatic
PA	Proto-Austric
PAA	Proto-Austro-Asiatic
PAN	Proto-Austronesian
PEAA	Proto-Eastern Austro-Asiatic
PMP	Proto-Malayo-polynesian

Note

1 Ross also faulted me for including a reference to Hayes’ basic vocabulary reconstructions, but not providing any examples. At the time when my paper was presented (at 8ICAL), Hayes’ paper was not yet published, and I had been given a pre-publication version of it, with the promise that I could refer to it but not cite any of his data. In addition, I had not at that time had the opportunity of evaluating the quality of his reconstructions.

Bibliography

- Adelaar, K.A. (1992) *Proto-Malayic: The Reconstruction of its Phonology and Parts of its Morphology and Lexicon*, Canberra: Pacific Linguistics C-119.
- Blust, R.A. (1993) ‘Central and Central-Eastern Malayo-Polynesian’, *Oceanic Linguistics* 32, 2: 241–94.
- (1999) ‘Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics’, in E. Zeitoun and P.J.-K. Li (eds) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, Taipei: Academia Sinica.

- Diffloth, G. (1994) 'The lexical evidence for Austric, so far', *Oceanic Linguistics* 33: 309–22.
- Hayes, L.H. (1997) 'On the track of Austric: Part II. Consonant mutation in early Austroasiatic', *Mon-Khmer Studies* 27: 13–44.
- (1999) 'On the track of Austric: Part III. Basic vocabulary comparison', *Mon-Khmer Studies* 29: 1–34.
- Parkin, R. (1991) *A Guide to Austroasiatic Speakers and Their Languages*, Oceanic Linguistics Special Publication, no. 23, Honolulu: University of Hawai'i Press.
- Reid, L.A. (1994) 'Morphological evidence for Austric', *Oceanic Linguistics* 33, 2: 323–44.
- (1999) 'New linguistic evidence for the Austric hypothesis', in E. Zeitoun and P.J.-K. Li (eds) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, Taipei: Academia Sinica.
- Ross, M.D. (2000) 'Review of E. Zeitoun and P.J.-K. Li (eds) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*', *Oceanic Linguistics* 39, 2: 445–56.
- Schmidt, W. (1906) 'Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens', *Archiv der Anthropologie, n.s.* 5: 59–109.
- (1916) 'Einiges über das Infix *mn* und dessen stellvertreter *p* in den austroasiatischen Sprachen', *Aufsätze zur Kultur- und Sprachgeschichte, vornehmlich des Orients, Ernst Kuhn zum 70*, Breslau: Marcus.
- Thurgood, G. (1999) *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change*, Oceanic Linguistics Special Publication No. 28, Honolulu: University of Hawai'i Press.

First published 2005
by RoutledgeCurzon
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by RoutledgeCurzon
270 Madison Ave, New York, NY 10016

RoutledgeCurzon is an imprint of the Taylor & Francis Group

© 2005 Editorial matter and selection, Laurent Sagart, Roger Blench and
Alicia Sanchez-Mazas; individual chapters, the contributors

Typeset in Times New Roman by
Newgen Imaging Systems (P) Ltd, Chennai, India
Printed and bound in Great Britain by
MPG Books Ltd, Bodmin

All rights reserved. No Part of this book may be reprinted or
reproduced or utilised in any form or by any electronic,
mechanical, or other means, now known or hereafter
invented, including photocopying and recording, or in any
information storage or retrieval system, without permission in
writing from the publishers.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record of this book has been requested

ISBN 0-415-32242-1