

[Warning: Draw object ignored]

The relatedness of Timor-Kisar and Alor-Pantar languages: A preliminary demonstration

Antoinette Schapper, Juliette Huber and Aone van Engelenhoven

The Papuan languages of Timor, Alor, Pantar and Kisar have long been thought to be members of a single family. However, their relatedness has not yet been established through the rigorous application of the comparative method. Recent historical work has shown the relatedness of the languages of Alor and Pantar on the one hand (Holton et al. 2012), and those of Timor and Kisar on the other (Schapper et al. 2012). In this chapter, we present a preliminary demonstration of the relatedness of the Timor-Alor-Pantar family based on a comparison of these two reconstructions. We identify a number of regular consonant correspondences across cognate vocabulary between the two groups and reconstruct a list of 89 proto-TAP roots.

1 Introduction

This chapter looks at the historical relationship between the Papuan languages of Alor-Pantar (AP) and those of Timor-Kisar (TK). The TK group of Papuan languages consists of Bunaq, spoken in central Timor, Makasae, Makalero and Fataluku, three languages spoken in a contiguous region of far eastern Timor, and Oirata, spoken on the southern side of Kisar Island to the north of Timor (Map 1). Due to their geographical proximity, AP and TK languages have typically been assumed to be related to one another (e.g., Stokhof 1975; Capell 1975). Together they have been referred to as the Timor-Alor-Pantar (TAP) family. However, there has been no substantive data-driven investigation of the claim.

In this chapter, we test the hypothesis that AP and TK languages are related to one another through the application of the comparative method. Specifically, we compare the results of two recent reconstructions, the one of AP (Holton et al. 2012) and the other of TK (Schapper et al. 2012). The sources of the lexical data used are listed in the Appendix. We establish that the AP and TK languages are indeed related by demonstrating that there are regular sound correspondences across cognate vocabulary between the two groups.

In comparing Holton et al. (2012) and Schapper et al. (2012) in this chapter, we assume the existence of two nodes in the TAP tree, namely Proto-AP (pAP) and Proto-TIM (pTIM). Whilst pAP appears to be a robust node, the existence of pTIM is less secure. As Schapper et al. (2012: 227-228) point out, it is possible that Bunaq and the Eastern Timor languages (reconstructed as Proto-ET in Schapper et al. 2012) both form their own separate primary subgroups within TAP. Our aim here is not to make claims about the high-level subgrouping of the AP and TK languages, and we do not presume to definitively determine the constituency of the TK-AP tree at this stage, but merely seek to show that TK and AP languages are related. Conclusive evidence of innovations shared by Bunaq and ET languages to the exclusion of AP languages is the subject of ongoing research.

Map 1: The Papuan languages of Timor and Kisar †

† Hatching marks areas where Papuan languages are found. Only Timor-Kisar languages are marked by name.

Section 2 presents the sound correspondences we find in cognate vocabulary between pAP and pTIM. Section 3 summarises our preliminary findings and discusses issues arising out of them. An Appendix is included with supporting language data for any reconstructions that do not appear in Holton et al. (2012) or Schapper et al. (2012), as well as a list of pTAP forms that can be reconstructed on the basis of the sound correspondences identified in this chapter. New, additional reconstructions have in some cases been necessary since the two articles each reconstruct only a small number of lexemes with only partial overlap between them. We also throw out several cognate sets from the AP reconstruction as they reflect borrowing from Austronesian languages.

2 Sound correspondences

In this section, we describe the consonant correspondences that we have identified between AP and TK languages. We do draw on vowel correspondences where they condition particular sound changes in consonants, but otherwise do not deal with vowels in this preliminary demonstration of relatedness. We chiefly draw attention to the correspondences in cognate vocabulary between pAP and pTIM. However, we provide the reader also with the forms of the lexemes in the TK languages as they are not available

elsewhere in this volume. The argumentation and underpinning data for pAP is given in Holton and Robinson (this volume) and is based on Holton et al. (2012).

In the subsections that follow, transcription of language data adheres to IPA conventions. Long vowels are indicated with a length mark ‘:’. Bracketed segments ‘()’ are those deemed to be non-etymological, that is, typically reflecting some morpheme which has fossilised on a root. In the correspondence tables, square brackets ‘[]’ are used where an item is cognate but doesn’t reflect the segment in question. The inverted question mark ‘?’ is used where a cognate shows an unexpected reflex of the segment in question. Grammatical items are glossed in small caps. Reconstructions marked with ‘!’ are new reconstructions not found in Holton et al. (2012) or Schapper et al. (2012), or are revised from Holton et al. (2012). The symbol ‘!’ signals that the full data set on which the reconstruction in question is based is given in the appendices. AP data supporting the additional pAP reconstructions is given in Appendix I and TK data in Appendix II. In the text of the chapter itself, for reasons of compactness, we only give simple one word glosses which reflect the presumed meaning of the proto-lexeme. Should the reader need more information, he can refer to the appendix. We also do not provide information on irregular changes, such as metathesis or apocope, in the correspondence tables, except where directly relevant to the reconstruction of the segment in question. The appendix provides the reader with fuller information on any irregularities in form or meaning in individual languages.

2.1 Reconstruction of bilabial stops

We identify two robust correspondent sets for bilabial plosives, reconstructing to pTAP *p and *b. Note that in Schapper et al. (2012), we reconstruct a three-way distinction (*p, *b, and *f) for bilabial obstruents in pTIM, despite the fact that it is not maintained in any of the modern TK languages: Bunaq, Makasae and Fataluku have merged reflexes of pTIM *p and *f, whereas in Fataluku and Oirata, *f and *b are merged. We find no evidence to support a three-way split in pTAP; instead, it looks like pTIM underwent a conditioned phoneme split, with distinct reflexes of pTAP *b in initial and non-initial positions, respectively.

Table 1 and Table 2 present the forms for these two correspondence sets respectively. In the first, pAP *p corresponds to pTIM *f in all positions. In the second, pTAP *b was retained as *b in pAP, but split to pTIM *b initially and pTIM *p non-initially. In these sets, there are three notable irregularities: (i) pAP *tiara ‘expel’ lost the medial bilabial that is retained in pTIM *tifar ‘run’; (ii) pAP *siba(r) ‘new’ and pTIM *(t,s)ifa(r) ‘new’ show an irregular correspondence of pAP *b with pTIM *f; and (iii) pAP *karab ‘scratch’ and pTIM *gabar ‘scratch’, which show an irregular correspondence of pAP *b with pTIM *b.

Table 1: Correspondence sets for pTAP *p

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *p	*p	*f	p, w	f	f	f	p
spit	*purVn !!	*fulu(k, n) !!	puluk	--	fulun	fulu	--
taboo	*palol !!	*falu(n)	por	falun	falun	falu	--
1PI	*pi-	*fi	--	fi	fi	afa	ap-
LOW ¹	*po !!	*ufe !!	--	he- ʔ	ufe-	[ua]	[ua]
girl	*pon !!	*fana ²	pana	fana(rae)	fana(r)	fana(r)	pana(rai)
scorpion	*pVr	*fe(r, R)e !!	wele	--	--	--	--
medial *p	*p	*f	w, Ø	f	f	f	p
face	*-pona !!	*-fanu !!	-ewen	fanu	fanu	fanu	panu
dream	*hipar	*ufar(ana) !!	waen	ufarena	ofarana	ufarana	upar(a)
run	[*tiara]	*tifar	tʃiwal	[ditar]	[titar]	tifar(e)	tipar(e)
pound	*tapai	*tafa	tao ³	--	tafa	tafa	tapa

¹ This item is a deictic marker for items at lower elevation than the deictic center. See Schapper (this volume) for more information on this deictic distinction.

² The bracketed *rae/r/rai* element appears to be an innovation in the Eastern Timor languages, presumably a lexical doublet or a derivational morpheme related to the nominalizing *-r* formative found in Makalero. We have no evidence for reconstructing this element higher than Proto-Eastern Timor.

³ This would have originally been *tawo in pre-Bunaq, but in the modern language medial /w/ is not preserved before back vowels.

Table 2: Correspondence sets for pTAP *b

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *b	*b	*b	b	b	p	p	h
pig	*baj	*baj	--	baj	paj	paj	haj
price	*bol !!	*bura	bol	bura	pura	pura	hura
mat	*bis	*biti !!	--	--	piti	pet(u)	het(e)
leg	*-bat !!	*-buta !!	-but	--	--	--	--
mountain	*buku !!	*bugu !!	--	bu?u	pu?u	--	--
non-initial *b	*b	*p	p, w	f	f	p	h
fish	*habi !!	*hapi !!	--	afi	afi	api	ahi
star	*jibV ¹	*ipi(-bere)	[bi] ²	ifi-bere	ifi	ipi(naka)	ihi
shark	*sib(a,i)r !! ³	*supor !!	--	--	[su] ⁴	hopor(u)	--
sugarcane	*hu:ba !!	*upa	up	ufa	ufa	upa	uha
tongue	*-lebur !!	*-ipul	-up	ifi	ifil	epul(u)	uhul(u)
dog	*jibar !! ⁵	*Depar	zap	defa	sefar	ipar(u)	ihar(a)
other	*aben(VC) !!	*epi !!	ewi	--	--	--	--
scratch	*karab !!	*gabar ɿ !! ⁶	--	--	kapar	kafur(e)	--
new	*siba(r) !!	*(t,s)ipa(r) ɿ !!	tip	sufa	hofar	--	--

¹ Several AP languages have a compound for ‘star’, although the second element does not appear to be cognate to that reconstructed for pTIM. Note also that Holton et al. (2012) gave this item as *jibC; this error has been corrected for Holton and Robinson (this volume).

² The Bunaq form reflects the second half of the pTIM doublet that is not found in AP languages.

³ The cognate set for this item is given in Holton et al. (2012), but no pAP reconstruction is given.

⁴ The reflex of the relevant bilabial has been lost in Makalero due to apocope.

⁵ The cognate set for this item is given in Holton et al. (2012), but no pAP reconstruction is given.

⁶ This form shows liquid-stop metathesis. There is no evidence of *b occurring word-finally in pTIM.

At this stage, we have no evidence for the reconstruction of a third bilabial obstruent to pTAP, as is found in pTIM (*p, *b and *f), but not pAP (*p and *b). Based on the current correspondence sets, the three-way distinction appears to have arisen due to pTAP medial *b changing to pTIM *p, while pTAP *b stayed *b initially in pTIM. We are yet to find any AP cognates for words reconstructing with initial *p in pTIM.

2.2 Reconstruction of coronal stops

There are two coronal stops, *t and *d, reconstructed to pAP, and four, *t, *d, *T and *D to pTIM. Schapper et al. (2012) note the uncertainty of pTIM *d, which is supported by three cognate sets only, all of which are in initial position. This is played out also when comparing coronals between AP and TK languages. We can reconstruct the pTAP coronal stops *t with relative certainty, and *d, albeit with less security. The latter segment split in pTIM to *T and *D. At present, we cannot reconstruct pTIM *d to pTAP. There are, however, a substantial number of coronal correspondences which remain unexplained.

Our most consistent correspondence is pTIM *t to pAP *t and *s (Table 3). Initially, we find a steady and unchanging correspondence of pAP *t and pTIM *t, supported by a sizeable number of cognates.

Only Bunaq shows a change of *t to /tʃ/ before a high front vowel. Non-initially, we find fewer cognates, but nevertheless a steady and unchanging correspondence. In two cognate sets (‘sit’ and ‘mat’), pAP final *s preceded by *i corresponds to pTIM *t.

Table 3: Correspondence sets for pTAP *t

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *t	*t	*t	t, tʃ	t	t	t	t
hand	*-tan	*-tana	-ton	tana	tana	tana	tana
sea	*tam	*mata	[mo]	--	--	mata	mata
six	*talam	*tamal !! ¹	tomol	--	--	--	--
pound	*tapai	*tafa	tao	--	tafa	tafa	tapa
run	*tiara	*tifar	tʃiwal	ditar ɿ	titar	tifar(e)	tipar(e)
sleep	*tia	*tia(r)	tʃier	taʔe	tia	taia	taja
non-initial *t	*t, *s	*t	t	t	t	t	t
tree	*tei	*hate !!	hotel	ate	ate	ete	ete
stand	*nate(r) !!	*nat	net	[na] ɿ	nat	(a)nat(e)	nat(e)
clew	*maita !!	*matar	mot	--	--	matar(u)	matar(a)
flat	*tatok !!	*tetok !!	toiʔ	--	tetuʔ	--	--
leg	*-bat !!	*-buta !!	-but	--	--	--	--
sit	*mis	*mit	mit	mit~[mi]	mit	[(i)mir(e)] ɿ	[mir(e)] ɿ
mat	*bis	*biti !!	--	--	piti	pet(u)	het(e)

¹ Bunaq /o/ is a regular reflex of pTIM *a, as seen, for instance, from the ‘hand’, ‘sea’, ‘wake’, ‘tree’ and ‘clew’ sets.

The reconstruction of pTAP *d is supported by only a small number of cognate sets (Table 4) and therefore still needs confirmation. In these sets, initial pAP *d corresponds to pTIM *D, while non-initial pAP *d corresponds with pTIM *T. This is consistent with what we observed with the bilabial stops, where a medial voiced stop in pAP corresponds to a voiceless stop in pTIM. Note that the cognate set for ‘bird’ is listed under the heading of initial *d, even though its pTIM and (arguably) pAP reflexes are in medial position. We place it there due to the fact that the sound correspondence is parallel to that for ‘rat’. However, more sets supporting this reconstruction are clearly needed before we can be certain of it.

Table 4: Correspondence sets for pTAP *d

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *d	*d	*D	z, s	d, s	s	c	t, s
rat	*dur	*Dura	zul	dura	sura	cura	ʈura
dog	*jibar ɿ !! ¹	*Depar	zap	defa	sefar	[ipar(u)] ɿ	[ihar(a)] ɿ
bird	*(a)dVl !!	*haDa	hos	asa	asa	aca	asa
medial *d	*d	*T	t	t	t	c	t
bat	*madel	*maTa !! ²	--	--	--	maca	maʈa
fire	*hada !!	*haTa	hoto	ata	ata	aca	aʈa
far	*lete !!	*eTar !!	ate	--	--	icar	--
sun	*wadi !!	*waTu	hot	watu	watu	wacu	waʈu
garden	*magad(a)	[(u, a)mar] ³	mar	ama	ama	--	uma

¹ We note the irregularity of pAP *jibar ‘dog’ where we would expect pAP *dipar ‘dog’. This is likely the result of a change pre-pAP *d > *j

² The cognate set for this item is given in Schapper et al. (2012), but no pTIM reconstruction is given.

³ This form shows metathesis with associated loss of the syllable with pTAP *g, thus: pTAP *magad > *madag > *amar.

Furthermore, there are a range of cognate sets which show as yet unexplained correspondences (Table 5). In these, we find coronal correspondences between pAP and pTIM and between TK languages (especially in Bunaq and Fataluku) that don’t fit well in the above given sets. More work is needed to clarify the history of the coronals in TAP.

Table 5: Problematic coronal cognate sets

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
grandparent ¹	*tam(a, u) !!	*moTo	mata(s)	mata	mata	moco	moŋo
far	*lete !!	*eTar !!	ate	--	--	icar	--
wake	*-ten	*Tani	otin	tane	tane	tani ~ cani	--
coconut	*wata !!	*wa(t, D)a	hoza	wata	wata	βata	wata
<i>P. indicus</i>	*matar !!	*ma(t, D)ar	mazo?	mater	mater	matar(ia)	--
excrement	*has	*a(t, D)u !!	ozo	atu(-gu?u)	atu	atu	atu

¹ This is a reciprocal kinship term, denoting either ‘grandparent’ or ‘grandchild’. PTIM *moTo means ‘child’.

2.3 Reconstruction of velar stops

We reconstruct two velar stops for pTAP, *k and *g. We find insufficient evidence, however, for the uvular stop reconstructed for pAP in Holton et al. (2012) and Holton and Robinson (this volume).

PTAP *k and *g are retained as *k and *g in pAP, but merged to *g in pTIM. Note that, based on the comparative TAP evidence and the additional pTAP reconstructions in this chapter, we have to substantially revise Schapper et al.’s (2012) pTIM reconstructions with regard to velar stops. Concretely, we can trace only one pTIM velar back to pTAP. We find no pAP reflexes for any of the small sets of roots reconstructed for pTIM with initial *k and medial *g; those for pTIM medial *g, in particular, are rather tenuous, as noted in Schapper et al. (2012: 212). The cognate sets that we can trace back to pTAP involve Schapper et al.’s initial *g and medial *k, and the comparative evidence is consistent with these being differential realisations of a single pTIM segment *g: initially, pTIM *g is reflected as /g/ in Bunaq and Makasae, and as /k/ in Makalero and Fataluku. We currently only have no evidence for Oirata. In non-initial position, *g is reflected in Bunaq as /g/ medially and as /k/ finally, consistent with Bunaq phonotactic rules, which prohibit voiced stops from codas; in Makasae, Makalero and Fataluku, *g is reflected in non-initial position as /ʔ/, and variably as /ʔ/ and Ø in Oirata.

The cognate sets that support the reconstruction of pTAP *k are given in Table 6.

Table 6: Correspondence set for pTAP *k

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*k	*g	g (k)	(g) ʔ	k, ʔ	k, ʔ ¹	(ʔ) Ø ¹
scratch	*karab !!	*gabar !! ²	--	--	kapar	kafur(e)	--
bite	*(ta)ki !! ³	*(ga)gel !! ³	gagil	gaʔel	kaʔel	(ki)kiʔ(e) ⁴	--
dirty	*karok !!	*gari !!	gar	raʔi	raʔi	raʔe(ne)	--
walk 1	*laka !!	*lagar !!	lagor	laʔa	laʔa	laʔa	[lare] ɿ
itchy	*(i)ruk !! ⁵	*ilag !!	--	ilaʔ	ileʔ	--	--
mountain	*buku !!	*bugu !!	--	buʔu	puʔu	--	--

¹ See Schapper et al. (2012: 211-212) for more Fataluku and Oirata correspondences.

² This form shows liquid-stop metathesis.

³ The bracketed initial segments in these forms reflect different inflectional prefixes which have fossilized on these verbs.

⁴ The initial bracketed syllable is a fossilized reduplicated CV. This item also has the variant pronunciation *ciki[241ʔ]e*.

⁵ This form represents a different root from the ‘itchy’ root given in Holton and Robinson (this volume). See Appendix I for supporting AP forms.

As in both pTIM (Schapper et al. 2012: 213-214) and pAP (Holton et al. 2012: 98), the reconstruction of initial *g in pTAP hinges on third person markers. Two forms are reconstructable (Table 7): a prefix

*g(a,i)- ‘3INLN’ occurring on verbs and inalienably possessed nouns, and a free form *gie ‘3ALN’ encoding 3rd person alienable possessors. Number marking was lost in TK languages, so the correspondence we observe is between pAP third person singular forms and pTIM third person forms which are unmarked for number (i.e., can be used in singular and plural contexts). The zero correspondence that we observe in Fataluku and Oirata is the result of the stripping off of the *g marking 3rd person (as set out in Schapper et al. 2012: 214). In the case of the alienable possessive marker, this means we are left with the possessive root pTIM *-ie ‘ALN’ alone.

Table 7: Correspondence set for pTAP 3rd person prefixes

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*g	*g	g	g	k	Ø	Ø
3INLN	*ga-	*g-	g-	g-	k-	--	--
3ALN	*ge ¹	*gie	gie	gi	ki	i	ue

¹ We reconstruct this as a free form on account of the existence of free reflexes in at least two AP languages (Blagar and Adang); morphologisation must thus post-date the break-up of pAP.

In non-initial positions, we find numerous cognates reflecting pTAP *g, corresponding to pAP *g and pTIM *g as set out in Table 8.

Table 8: Correspondence set for pTAP *g

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*g	*g	g, k	g, ?	(k) ?	(k) ?	?, Ø
yellow	*bagori !! ¹	*gabar !! ²	--	gabar	--	--	--
green	*(wa)logar !!	*ugar	ugar	(h)u?ur	(h)u?ur	u?ur(eke)	u?ul(e)
laugh	*jagir !!	*jiger !!	higal	hi?a	hi?e	he?e	--
path	*jega !!	*jiga !!	hik	hi?a	hi?a	i?a	ia(ra)
banana	*mogol	*mugu !!	mok	mu?u	mu?u	mu?u	mu:
hear	*magi !! ³	*mage(n) !!	mak	ma?en	ma?en	--	--
garden	*magad(a)	[*(u, a)mar] !! ⁴	mar	ama	ama	--	uma

¹ The cognate set for this item is given in Holton et al. (2012), but no pAP reconstruction is given.

² This form is apparently metathesized from pTAP *bagori ‘yellow’.

³ The cognate set for this item is given in Holton et al. (2012), but no pAP reconstruction is given.

⁴ This form shows metathesis with associated loss of the syllable with pTAP *g, thus: pTAP *magad > *madag > *amar. Loss of *g is found occasionally in AP languages (e.g. ‘laugh’, see Appendix 1), suggesting a certain degree of instability for this segment.

Finally, there is as yet an insufficient number of reconstructions of pAP *q with cognates in TK languages to allow for a higher-level pTAP reconstruction. Currently, we have only Bunaq -ol ‘child’ (presumably reflecting pTIM *-al) as cognate with pAP *-uaqal ‘child’. We await further reconstructions with TK cognates for the determination of the pTAP form.

2.4 Reconstruction of fricatives

Two fricatives *s and *h can be reconstructed to pTAP. The number of cognates is still small for both phonemes, but the correspondences are relatively well-behaved.

Table 9 sets out the cognate sets for pTAP *s. Initial pTAP *s is supported by several cognate sets and has been maintained without change in pAP and pTIM. Non-initial cognates of pAP *s are difficult to find in TK languages, as many instances of reconstructed word-final *s in pAP correspond to pTIM *t (e.g., pAP *mis ‘sit’, *bis ‘mat’ and *has ‘excrement’ [336?] see also our discussion of these in section 4).

Table 9: Correspondences of pTAP *s

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata ¹
initial *s	*s	*s	s	s	h, s ²	h	s
bone	*ser !!	*(se)sa(r, R) !!	sesal	--	--	--	--
shark	*sib(a, i)r !!	*supor !!	--	--	su-	hopor(u)	--
spoon	*surV !!	*sula	sulu	sulu	hulu	hula	sulu
weave	*sine(N) !!	*sina	sien	sina	hina	hina	hina(na) ɿ
new	*siba(r) !! ³	*(t, s)ipa(r) !! ₃	tip ɿ	sufa	hofar	--	--
non-initial *s	*s	*s	s	s	s	h	Ø
meat	*iser !! ⁴	*seor	saɛl	seu	seur	[leura] ɿ	[leura] ɿ
tooth	*-uasin !!	*-wasin !!	[(-e)we] ɿ	wasi	wasi	ʃahin(u)	wain(i)

¹ See Schapper et al. (2012: 209) for more instances of Oirata cognates.

² Makalero seems to be part-way through a sound change s > h. See Schapper et al. (2012: 209-211) for more cognates showing the variable s ~ h reflexes in Makalero.

³ Cognates for these reconstructions show a relatively high degree of irregularity in both AP and TK indicating that there may have been variable realisations in not only pAP and pTIM, but also pTAP.

⁴ Denotes ‘meat’ or ‘game’

PTAP *h can be reconstructed as a word-initial segment, but not in other positions. The segment corresponds to pTIM *h and pAP *h except before back vowels (Table 10). Based on the cognate sets available, pAP *h did not occur before back vowels. In this environment, pTAP *h changed either to *w (as in pAP *wur ‘moon’) or was lost (as in pAP *tei ‘tree’) in pAP (cf. Table 11 for the items and vocalic environments in which pAP *w is attested). The reconstruction pTIM *h hinges on Bunaq, which retains it as /h/, while the eastern Timor languages have all lost pTIM *h (which, in turn, reflects pTAP *h). This means that where we have no Bunaq reflex (as in the ‘fish’, ‘breast’ and ‘dream’ sets) we have no modern language attesting pTIM *h, and the presence of the phoneme can only be inferred from the fact that *h is reconstructed for the pAP cognate.

Table 10: Correspondence set for *h

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*h (*w/Ø)	*h	h	Ø	Ø	Ø	Ø
fire	*hada !!	*haTa	hoto	ata	ata	aca	aʔa
fish	*habi !!	*hapi !!	--	afi	afi	api	ahi
breast	*hami	*hami !!	--	ami	--	ami(-tapunu)	--
moon	*wur	*huru	hul	uru	uru	uru	uru
tree	*tei ¹	*hate !!	hotel	ate	ate	ete	ete

¹ The loss of initial syllable may have to do with the fact that stress was apparently based on syllable weight. See also ‘dog’ in Appendix 1 and Holton and Robinson (this volume).

2.5 Reconstruction of glides

Two glides can be reconstructed to pTAP, *w and *j. Both appear to have only occurred in initial position. It is unclear whether the reconstructed glides could occur before all vowel qualities. Nevertheless, the cognate sets supporting these proto-phonemes are robust and show little irregularity.

The pTAP glide *w shows a stable and unchanging correspondence of *w in pAP and pTIM for the most part (Table 11). The major change is that pTAP *w is vocalised in pAP to *u root-initially on inalienably possessed nouns. In TK languages, Bunaq shows conditioned reflexes of pTAP *w, maintaining it as /w/ before front vowels, but changing it to /h/ before non-back vowels. Fataluku shows the change of *w to /ʃ/, though we note that this is an allophone of /w/ in many languages.

Table 11: Correspondence set for pTAP *w

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*w, *u	*w	h, w	w	w	ʃ	w
blood	*wai	*waj	ho	waj	wej	ʃehe	we
coconut	*wata !!	*wa(t, D)a	hoza	wata	wata	ʃaca	wata
stone	*war	*war	hol	--	war	--	war(aha)
sun	*wadi !!	*waTu	hot	watu	watu	ʃacu	waʃu
bathe	*weli	*weru	wer	waru?	waro?	ʃahu	wau
ear	*-uari !!	*-wali	--	wala(ku)	wali	ʃali	wali
tooth	*-uasin !!	*-wasin !!	-(e)we	wasi	wasi	ʃahin(u)	wain(i)

Table 12 gives the four clear cognate sets that we have across TAP languages for pTAP *j. We see that pTAP *j is maintained as *j in pAP, but is variably lost or maintained as *j in pTIM. It may be that differing vocalic environments in pTAP conditioned the different reflexes in pTIM, but we don't have enough understanding of the history of vowels yet to determine this. There is no direct evidence for pTIM *j, that is, no TK language still reflects the proto-phoneme as /j/, but the sound correspondences between TK languages make it differentiable form sets reflecting pTIM *h (see Table 10).

Table 12: Correspondence set for pTAP *j

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*j	j, Ø	h, Ø	h, Ø	h, Ø	Ø	Ø
star	*jibV	*ipi(-bere)	[bi]	ifi-bere	ifi	ipi(-naka)	ihi
water	*jira	*ira	il	ira	ira	ira	ira
laugh	*jagir !!	*jiger !!	higal	hiʔa	hiʔe	heʔe ɿ	--
path	*jega !!	*jiga !!	hik	hiʔa	hiʔa	iʔa	ia(ra)

2.6 Reconstruction of liquids

We identify three robust liquid correspondence sets between pAP and pTIM and as such reconstruct three pTAP liquids: *r, *R, and *l.

The most robust set is that for pTAP *r, which is reflected as *r in both pAP and pTIM (Table 13). PTAP *r is only found in non-initial positions, as are its reflexes in the daughter languages pAP and pTK. Word-finally in polysyllabic words pTAP *r is particularly susceptible to sporadic loss, as is attested by the various irregular forms in Table 13. In one instance (pTAP *sibar 'new'), the occurrence of a reflex of final *r is so erratic in both primary subgroups that we perhaps must consider it already partly lost in pTAP's daughter languages.

Table 13: Correspondence set for pTAP *r

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*r	*r	l	r	r	r	r
run	*tiara	*tifar	tʃiwal	ditar	titar	tifar(e)	tipar(e)
moon	*wur	*huru	hul	uru	uru	uru	uru
rat	*dur	*Dura	zul	dura	sura	cura	ʈura
stone	*war	*war	hol	--	war	--	war(aha)
vagina	*-ar	*-aru	--	aru	aru	aru	aru
water	*jira	*ira	il	ira	ira	ira	ira
crawl	*er !!	*er !!	el	--	--	er(eke)	--
dream	*hipar	*ufar(ana) !!	[waen] ɿ	ufarena	ofarana	ufar(e)	upar(a)
meat	*iser !!	*seor	sacl	[seu] ɿ	seur	leura	leura
dog	*jibar !!	*Depar	[zap] ɿ	[defa] ɿ	sefar	ipar(u)	ihar(a)
bamboo	*mari	*mari	[ma] ɿ	maeri	mar	--	--
<i>P. indicus</i>	*matar !!	*ma(t, D)er	[mazoʔ] ɿ	mater	mater	matar(ia)	--
shark	*sib(a, i)r !!	*supor !!	--	--	[su] ɿ	hopor(u)	--
new	*siba(r) !!	*(t, s)ipa(r) !!	[tip] ɿ	[sufa] ɿ	hofar	--	--

PTAP *R is reflected in pAP as *r and in pTIM as *l. Like pTAP *r, *R does not appear in word-initial positions and is sporadically lost word-finally in polysyllabic words. The sets supporting the reconstruction of *R (Table 14) are also fewer and less robust than for pTAP *r.

Table 14: Correspondence set for pTAP *R

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*r	*l	l	l	l	l	l
spoon	*surV !!	*sula	sulu	sulu	hulu	hula	--
tail	*-ora !!	*-ula(?)	-ulo(?)	ula	ula	ula(fuka)	ula(pua)
tongue	*-lebur !!	*-ipul	[-up] ɿ	[ifi] ɿ	ifil	epul(u)	uhul(u)
laugh	*jagir !!	*jiger	higal	[hiʔa] ɿ	[hiʔa] ɿ	[heʔe] ɿ	--
spit	*purVn !!	*fulu(k, n) !!	puluk	--	fulun	fulun	--
ear	*-uari !!	*-wali	--	wala(ku:)	wali	ʃali	wali

The three pTIM cognates listed in Table 15 are based on Bunaq only, in which pTIM *r and *R are merged. We have thus no means of determining whether these forms are to be reconstructed to pTAP with *r or with *R.

Table 15: Cognate sets reconstructable to either pTAP *r or *R

	pAP	PTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*r	*(r, R)	l	--	--	--	--
bone	*ser !!	*(se)sa(r, R) !!	sesal	--	--	--	--
scorpion	*pVr	*fe(r, R)e !!	wele	--	--	--	--
rain	*anur !!	*ine(r, R) !!	inel	--	--	--	--

Cognate sets for pTAP *l are relatively infrequent in both pAP and pTIM (Table 16).¹ Cognates reflecting initial pTAP *l with pAP *l and pTIM *l (i.e., ‘bark’, ‘new place’ and ‘crouch’) have only a low degree of certainty. Based on the data available, there also appears to be a tendency to lose pTAP initial *l in pTIM, as in ‘far’, ‘tongue’ and ‘green’, but a clear conditioning environment for this is not yet obvious. Word-finally in polysyllabic words, pTAP *l is regularly lost in pTIM, as in ‘banana’, ‘bat’, ‘bird’ and ‘taboo’. However, it is retained in ‘walk 2’ and ‘six’, apparently due to nasal-liquid metathesis, and in ‘child’ due to the loss of the item’s medial syllable with *q prior to the application of the final polysyllabic deletion rule in pAP.

Table 16: Correspondence set for pTAP *l

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
	*l	*l, Ø	l (Ø)	l (Ø)	l (Ø)	l (Ø)	l (Ø)
bark	*lVu	*le(k)u(l) !!	--	leu	leu	leʔul(e)	leul(e)
new place	*lan !!	*lan !!	lon	--	--	--	--
crouch	*luk(V)	*luk !!	luʔ(-luʔ)	--	--	--	--
far	*lete !!	[*eTar] !!	ate	--	--	icar	--
tongue	*-lebur !!	[*-ipul]	-up	ifi	ifil	epul(u)	uhul(u)
green	*(wa)logar !!	[*ugar]	ugar	huʔur	(h)uʔur	uʔur(eke)	uʔul(e)
banana	*mogol	[*mugu] !!	mok	muʔu	muʔu	muʔu	mu:
bat	*madel	[*maTa] !!	--	--	--	maca	maʔa
bird	*(a)dVl !!	[*haDa]	hos	asa	asa	aca	asa
taboo	*palol !!	[*falu(n)]	por	falun	falun	falu	--
walk 2	*lam(ar) !!	*male !!	mele	--	--	--	--
six	*talam	*tamal !!	tomol	--	--	--	--
child	*-uaqal	*-al !!	-ol	--	--	--	--

Finally, there are several cases in which the appearance of liquids in AP and TK languages can be reconciled with none of the three sets we have identified here. Table 17 lists these problematic instances (the relevant segments are bolded). These sets pointedly express that we are still a long way away from a complete understanding of liquids in pTAP.

Table 17: Problematic liquid cognate sets

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
price	*bol !!	*bura	bol	bura	pura	pura	hura
bathe	*weli	*weru	wer	waruʔ	waroʔ	vahu ʔ	wau ʔ
garden	*magad(a)	*(u, a)mar	mar	[ama]	[ama]	--	[uma]
green	*(wa)logar !!	*ugar	ugar	(h)uʔur	(h)uʔur	uʔur(eke)	uʔul(e) ʔ
Taboo	*palol !!	*falu(n)	por	falun	falun	falu	--

2.7 Reconstruction of nasals

Two nasals can be reconstructed to pTAP, *m and *n. For the most part, they are relatively stable and unchanging in both pAP and pTIM.

Table 18 presents a selection of the many cognate sets for pTAP *m. In word-initial position, pTAP *m corresponds unproblematically to pAP *m and pTIM *m. Identifying non-initial instances of pTAP

¹ Holton and Robinson (this volume) remark that, even though correspondences appear relatively regular for initial and medial *l in pAP, they can identify only a few cognates that are widely distributed across the AP subgroup. Similarly, Schapper et al. (2012) caution that their reconstruction for pTIM *l cannot yet be called secure due to the small number of cognate sets identified.

*m is somewhat more difficult, with *hami ‘breast’ being the only straightforward case. Word-final *m in pAP has only non-final reflexes in pTIM, apparently because, as in the modern TK languages, word-final *m was not permitted. This issue is resolved in pTIM through metathesis of the nasal out of the final position, as in ‘sea’ and ‘six’. Other instances of medial pTIM *m correspond to root-initial *m in pAP (as in ‘garden’ and ‘die’).

Table 18: Correspondence sets for pTAP *m

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *m	*m	*m	m	m	m	m	m
bamboo	*mari	*mari	ma	maeri	mar	--	--
banana	*mogol	*mugu !!	mok	muʔu	muʔu	muʔu	mu:
sit	*mis	*mit	mit	mit ~ mi	mit	(i)mir(e)	mir(e)
bat	*madel	*maTa !!	--	--	--	maca	maʔa
inside	*mi	*mi	mi(l)	mu(tu)	mu(tu-)	mu(cu)	mu(tu)
hear	*magi !!	*mage(n) !!	mak	maʔen	maʔen	--	--
non-initial *m	*m	*m	m	m	m	m	m
breast	*hami	*hami !!	--	ami	--	ami(-tapunu)	--
sea	*tam	*mata	mo	--	--	mata	mata
six	*talam	*tamal !!	tomol	--	--	--	--
garden	*magad(a)	*(u, a)mar !!	mar	ama	ama	--	uma
die	*min(a)	*-umV	-ume	umu	(k)umu	umu	umu
nose	*-mim	*-muni !!	[-inup] ɿ	muni(kai)	mini	mini(ku)	--

Table 19 presents the many cognate sets for pTAP *n. Initial and medial correspondences are abundant, but final correspondences are difficult to identify. pTIM *n did not appear to occur in final position; all instances of pAP final *n are either followed by a vowel or are lost in pTIM.

Table 19: Correspondence sets for *n

	pAP	pTIM	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *n	*n	*n	n	n	n	n	n
stand	*nate(r) !!	*nat	net	nat ~ na	nat	(a)nat(e)	nat(e)
1sg	*na-	*n- !!	n-	--	--	--	--
eat	*nai	*nua !!	[a ~ -ia]	nawa	nua	una, naβa	una, nawa
one	*nuk	*uneki !!	uen, en	[u]	[u] ~ un	ukani	aʔuni
non-initial *n	*n	*n	n	n	n	n	n
face	*-pona !!	*-fanu !!	-(e)wen	fanu	fanu	fanu	panu
ripe	*tena !!	*tena !!	ten	tina	tina	--	--
name	*-en(i, u) !!	*-nej	-ini(l)	naj	nej	ne	ne:(ne)
give	*-ena	*-inV	-ini	(g)ini	(k-)ini	ina	ina
wake	*-ten	*Tani	otin	tane	tane	tani ~ cani	--
girl	*pon !!	*fana	pana	fana(rae)	fana(r)	fana(r)	pana(rai)
person	*anin !!	*anu !!	en	anu	anu	--	--
other	*aben(VC) !!	*epi !!	[ewi]	--	--	--	--

3 Summary of correspondences and reconstructed phonemes

For the first time since the start of TAP studies some sixty years ago (see Schapper and Huber forthcoming for a historical perspective on TAP studies), we have rigorously shown in this chapter that the TAP languages form a family: the regularity of sound correspondences in cognate vocabulary demonstrates that the AP and TK Papuan languages are indeed genetically related to one another.

In Table 20, we provide an overview of the consonant correspondences we observed in cognate vocabulary between pAP and pTIM and their reconstruction in their ancestral language pTAP. In this table, we indicate whether the correspondence applies in initial (# _), medial (V _ V), or final (_ #) position. An empty slot means that there is no particular conditioning environment for the correspondence. The symbol ‘Ø’ in a column indicates that a pTAP sound is lost in the daughter language in question.

Table 20: Summary of sound correspondences from pTAP to pAP and pTIM

pTAP	environment	pAP	pTIM
*p		*p	*f
*b	# _ V _ V	*b	*b
*t	# _ V _ V, _ #	*t, *s	*t
*d	# _ V _ V	*d	*D
*k		*k	*k
*g		*g	*g
*s		*s	*s
*h		*h (*w/Ø)	*h
*w		*w, *u	*w
*j		*j	*j, Ø
*r		*r	*r
*R		*r	*l
*l		*l	*l, Ø
*m		*m	*m
*n		*n	*n

4 Discussion

Whilst we have been able to show clearly that AP and TK languages are related to one another, the comparative data presented here draws into question a number of aspects of the existing reconstructions of pAP and pTIM and necessitates revisions to these. In this final section, we will draw attention to the issues, provide a general discussion of them and suggest some possible solutions to them.

A major issue for the current pAP reconstruction is the apparent invalidity of many word- final consonant reconstructions. It is argued in Holton et al. (2012: 95) that the gemination of medial stops in modern Western Pantar can be used as a diagnostic for determining whether a given pAP root was consonant-final or vowel-final. Specifically, the authors claim that geminate medial stops in modern Western Pantar reflect pAP medial stops, whereas non-geminate medial stops in Western Pantar reflect an original consonant-final form, or perhaps a borrowing from another AP language. However, this argument cannot be sustained on closer inspection of the comparative evidence. Consider the items in Table 21 that are reconstructed as basically consonant final in pAP, because of the lack of stop gemination in WP. In each case, we have between three and nine reflexes in modern AP languages with a V(C) following the supposed historically final consonant. We must ask ourselves where so many additional final segments came from in so many of these languages. Holton et al. (2012) seek to explain these appearances with

vowel epenthesis. Yet, under this scenario, we would expect to be able to predict the type of the epenthetic vowel from the shape of the root, but this is not the case; instead, the epenthetic vowels are of all different values from one item to the next and bear no apparent relationship to the vowel of the root (as defined by Holton et al. 2012). What is more, the final V(C) elements we observe in AP languages are not erratic, rather they in general adhere to correspondences observed elsewhere. This suggests that these final V(C) elements were not epenthetic to the items after the break-up of pAP, but have been inherited from pAP. This is further supported by the fact that we find clearly corresponding V(C) segments on cognate vocabulary in TK languages, meaning that the segments reconstruct to pTAP and that they were inherited into pAP. The alternative leaves us without explanation for the cognacy of the final segments in these (and other items) across the family.

Table 21: Dubious consonant-final reconstructions in AP and beyond

	‘fish’	‘sun’	‘fire’	‘coconut’	‘tongue’	‘ripe’
pTAP	*habi	*wad(u, i)	*hadi	*wata	*(l)ebur	*tena
pTIM	*hapi	*waTu	*haTa	*wa(t, D)a	*-ipul	*tena
pAP original	*hab(i)	*wad(i)	*had(a)	*wat(a)	*-leb(ur)	*ten
Teiwa	[127?]aɸ	war	[127?]ar	wat	-livi	tanana
Nedebang	a:fi	(get)	ar	wata	-lefu	--
Kaera	ab	wer	ad	wat	-leb	ten
WPantar	hap	wer	a:d	wata	-lebu	taŋ
Blagar	a:b	war	--	vet	-lebul	tena
Adang	a:b	ved		faʔ	-lib(uŋ)	tene
Klon	əbi	fəd	əda	--	-ləb	ətən
Kui	eb	--	ar	bat	-liber	tain
Abui	afu	wari	ara	wata	-lifi	--
Kamang	api	wati	ati	wate	-opui ¹	iten ~ iton
Sawila	api	wadi	ada	wata	-li(m)puru	itima
Wersing	api	widi	ada	wata	-jebur	--

¹ Holton et al. (2012) state that these and other Kamang forms missing pAP *l medially are irregular. However, pAP *l is regularly lost in Kamang between non-front vowels, e.g., pAP *talam ‘six’ > Kamang *ta:m*, pAP *palol ‘taboo’ > Kamang *fo:ɪ* etc. The vowel of the inalienable possessive prefix is /a/, thus providing the right environment for the loss in *-opui* ‘tongue’ of the root-initial /l/.

The problem then is how to explain medial geminate and non-geminate stops in WP. One answer would be to maintain that the difference in stop gemination was still due to a final- non-final distinction. For example, it could be said that the loss of the final vowel occurred after the breakup of pAP but prior to the application of the gemination rule. This cannot, however, be fully sustained as WP has in some cases final vowels which clearly reflect pTAP and pAP (e.g., ‘tongue’). A more attractive explanation is presented by stress-induced gemination. Although little is known about the historical prosody of TAP, it seems a good possibility that WP gemination may have been a result of final stress. That is, we suggest roots of the shape /-(C)V'CV(C)/ surfaced as [(C)V'C:V(C)], while roots of the shape /'-(C)V'CV(C)/ surfaced as [-(C)V'CV(C)]. While this scenario remains to be confirmed by a more detailed study, discarding the final/non-final explanation for geminates in WP allows for a more satisfactory account of final segments in TAP.

A second issue for the pAP reconstruction is the presence of many unexplained phonemes in a range of environments in different languages. Velars, post-velar and laryngeal consonants are a case in point. Most of the complexity in this domain is found in the languages of Pantar and the Pantar Straits, whose phoneme inventories generally include not only velar and glottal stops, but also uvular ones, as well as a velar or pharyngeal fricative next to the glottal fricative /h/. This contrasts with the situation as found in most of Alor and the TK languages, which tend to be rather simpler. Table 22 exemplifies the velar and post-velar plosives and fricatives in a language of Pantar (Teiwa), Alor (Kamang), and Timor (Bunaq).

Table 22: Velar and post-velar phonemes in TAP languages

	Teiwa				Kamang		Bunaq
	velar	uvular	pharyngeal	glottal	velar	glottal	velar
plosive	k g	q		ʔ	k g	(ʔ)	k g
fricative			[127ʔ]	h		(h)	

The existing pAP reconstruction leaves a significant part of the complexity in the (post)velar domain in the Pantar languages unexplained; for instance, it does not account for /g/ in Blagar and the relation between the various (post-)velar phonemes such as /q/ and /x/ found in different dialects of Blagar.² It also does not explain the origin of /ʔ/ in languages other than Blagar and Adang, and does not give reflexes for pAP medial *k in Teiwa and pAP final *k in Sawila, leaving the field in question blank in the table summarising the correspondences (Holton and Robinson this volume). Finally, note a variety of irregularities in the reconstructions involving velars in Appendix 1, especially in the Pantar languages. In short, the frequency of irregularities and unexplained occurrences of (post-)velar phonemes shows how limited our understanding of this domain in AP still is, and serves as a reminder that much more extensive reconstruction work needs to be undertaken.

A similar issue is presented by the phonemic velar nasal /ŋ/ in many AP languages. This phoneme is not reconstructed for pAP, and is also absent in all of the TK languages. According to Holton and Robinson (this volume), pAP *n became /ŋ/ in word-final position in all AP languages except Teiwa, where it was retained as /n/. This historical scenario does work well for some languages, for instance, Wersing, where [ŋ] is synchronically a word-final allophone of /n/. However, in other languages, questions remain. For instance, Kamang has an unexplained contrast between /ŋ/ and /n/ in codas (e.g., *een* ‘2SG.POSS’ versus *een* ‘2SG.FOC’). Similarly, the existence of /ŋ/ in coda and medial position in Teiwa is unexplained, as well as the occurrence of /ŋ/ in other positions than the final one in various languages (e.g. Sar *lanja* ‘digging stick’ and Kula *ŋapa* ‘father’).

Vowels also present a major challenge to the reconstruction of the ancestral TAP language. The various vowel systems as illustrated in Table 23 are yet to be historically reconciled with one another. Most AP languages have a length distinction in their vowels: the most common system is 5 short and 5 long cardinal vowels (Kaera, Blagar, Abui and Kamang), though matching long vowels may be missing in the mid-vowel range (Teiwa and Klon). Blagar has a marginal length distinction with only a small number of items occurring with long vowels (Steinhauer forthcoming), while it is Klon’s short mid-vowels that are marginal. A length distinction is entirely absent from WP’s and Wersing’s five vowel system and Adang’s seven vowel system. A relationship, if any, between the mid-vowels in Adang and length distinctions in other languages remains to be established. Non-cardinal vowels are found in Sawila /y, y:/ and in Klon /ə/. TK languages all have simple five cardinal vowels and there is a marginal length distinction in only one language, Makalero. Stress in conjunction with length appears to have played an important role in vowel histories. For instance, Klon /ə/ seems to originate in a short, unstressed pAP *a (e.g., Klon *əbi* appears to go back to pAP *ha’bi ‘fish’). In Wersing, historically short unstressed vowels are lost in words with long vowels, which in turn become short stressed vowels (e.g., Wersing *tlam* appears to go back to pAP *tala:m ‘six’, cf. Abui *tala:ma*). In short, much careful bottom-up reconstructive work needs to be done in order to reconcile these different systems to a single ancestral system.

Table 23: TAP vowel systems³

² See Steinhauer (1995).

³ The data in these tables are from Holton (forthcoming) for Western Pantar, Klammer (2010) for Teiwa, Klammer (forthcoming) for Kaera, Steinhauer (forthcoming) for Blagar, Haan (2001) for Adang, Baird (2008) for Klon, Kratochvil (2007) for Abui, Schapper (n.d.) for Kamang, Kratochvil (forthcoming) for Sawila, Schapper and Hendery (forthcoming) for Wersing, Schapper (2010) for Bunaq, and Huber (2011) for Makalero.

WP		Teiwa		Kaera		Blagar
i	u	i, i:	u, u:	i, i:	u, u:	i, i:
e	o	e	o	e, e:	o, o:	e, e:
a		a, a:		a, a:		a, a:
Adang		Klon		Abui		Kamang
i	u	i, i:	u, u:	i, i:	u, u:	i, i:

e	o	e	o, o:	e, e:	o, o:	e, e:
ɛ	ɔ	ɛ, ɛ:	ə	ɔ		
a		a, a:		a, a:		a, a:
Sawila		Wersing		Bunaq		Makalero
i, i: y, y:	u, u:	i	u	i	u	i
e, e:	o, o:	e	o	e	o	e
a, a:		a		a		a

In sum, with the positive establishment of the relatedness of the Papuan languages scattered across the islands of Timor, Kisar, Alor, Pantar and the Pantar Straits, a start has been made towards a history of the TAP languages. However, we are still a long way off a complete and nuanced understanding of the family and its development (cf. Schapper and Huber’s (forthcoming) statement of prospective research questions). It will be the task of future reconstructive historical work to definitively solve remaining issues in the comparative data.

5 References

- Baird, Louise. 2008. *A grammar of Klon: A non-Austronesian language of Alor, Indonesia*. Canberra: Pacific Linguistics.
- Brotherson, Anna. 2003. A spatial odyssey: Referring to space in Makasai. BA thesis, Australian National University, Canberra.
- Capell, Arthur. 1975. The West Papuan Phylum: General, and Timor and areas further west. In Stephen A. Wurm, ed., *New Guinea area languages and language study*, 667-716. Canberra: Pacific Linguistics.
- Carr, Felicita M. 2004. Pupuuk, Masu Moru and Data: Three genres of Makasai. MA thesis, The Australian National University, Canberra.
- Haan, Johnson. 2001. The grammar of Adang, a Papuan language spoken on the island of Alor, East Nusa Tenggara, Indonesia. PhD thesis, University of Sydney.
- Holton, Gary. Forthcoming. Western Pantar. In Antoinette Schapper, ed., *Papuan languages of Timor-Alor-Pantar: Sketch grammars Volume I*. Berlin/New York: Mouton de Gruyter.
- Holton, Gary, Marian Klamer, František Kratochvíl, Laura C. Robinson, and Antoinette Schapper. 2012. The historical relations of the Papuan languages of Alor and Pantar. *Oceanic Linguistics* 51 (1): 86-122.
- Holton, Gary, and Mahalalel Lamma Koly. 2008. *Kamus pengantar bahasa Pantar Barat*. Kupang, Indonesia: Unit Bahasa dan Budaya (UBB).
- Holton, Gary, and Laura C. Robinson. This volume. The internal history of the Alor-Pantar language family. In Marian Klamer, ed., *Alor-Pantar languages: History and typology*. Leiden: Brill.
- Huber, Juliette. 2008. *First steps towards a grammar of Makasae*. München: Lincom.
- Huber, Juliette. 2011. *A grammar of Makalero: A Papuan language of East Timor*. Utrecht: LOT.
- Josselin de Jong, J.P.B. de. 1937. *Studies in Indonesian culture I. Oirata, a Timorese settlement on Kisar*. Amsterdam: N.V. Noord-Hollandsche Uitgevers-Maatschappij.
- Klamer, Marian. 2010. *A grammar of Teiwa*. Berlin/New York: Mouton de Gruyter.
- Klamer, Marian. Forthcoming. Kaera. In Antoinette Schapper, ed., *Papuan languages of Timor-Alor-Pantar: Sketch grammars Volume I*. Berlin/New York: Mouton de Gruyter.
- Klamer, Marian, and Amos Sir. 2011. *Kosakata bahasa Teiwa-Indonesia-Inggris* [Teiwa-Indonesian-English word list]. Kupang, Indonesia: Unit Bahasa dan Budaya (UBB).
- Klamer, Marian, and Antoinette Schapper. 2012. The history of ‘give’ constructions in the Papuan languages of Timor-Alor-Pantar. *Linguistic Discovery* 10 (3): 174-207.
- Kratochvíl, František. 2007. *A grammar of Abui: A Papuan language of Alor*. Utrecht: LOT.
- Kratochvíl, František. Forthcoming. Sawila. In Antoinette Schapper, ed., *Papuan languages of Timor-Alor-Pantar: Sketch grammars Volume II*. Berlin/New York: Mouton de Gruyter.
- Kratochvíl, František. n.d. Dictionary of Sawila. [Manuscript, Nanyang Technological University.]

- Kratochvíl, František, and Benediktus Delpada. 2008. *Abui stories from Takalelang*. Kupang, Indonesia: Unit Bahasa dan Budaya (UBB).
- Rodemeier, Susanne. 1992. *Lego-lego Platz und naga-Darstellung. Jenseitige Kräfte im Zentrum einer Quellenstudie über die ostindonesische Insel Alor*. MA thesis, Ludwig-Maximilians-Universität München.
- Schapper, Antoinette. This volume. Elevational systems in AP languages. In Marian Klammer, ed., *Alor-Pantar languages: History and typology*. Leiden: Brill
- Schapper, Antoinette. 2010. *Bunaq, a Papuan language of central Timor*. PhD thesis, The Australian National University, Canberra.
- Schapper, Antoinette. n.d. a. Dictionary of Bunaq. [Manuscript, Australian National University.]
- Schapper, Antoinette. n.d. b. Grammar of Kamang. [Manuscript, Leiden University.]
- Schapper, Antoinette, and Rachel Hendery. Forthcoming. Wersing. In Antoinette Schapper, ed., *Papuan languages of Timor-Alor-Pantar: Sketch grammars Volume I*. Berlin/New York: Mouton de Gruyter.
- Schapper, Antoinette, and Juliette Huber. Forthcoming. State-of-the-art in the documentation of the Papuan languages of Timor, Alor, Pantar and Kisar. *Wacana, Journal of the Humanities of Indonesia*.
- Schapper, Antoinette, Juliette Huber, and Aone van Engelenhoven. 2012. The historical relations of the Papuan languages of Timor and Kisar. *Language and Linguistics in Melanesia: Special Issue on the History, Contact & Classification of Papuan languages*: 194-242.
- Schapper, Antoinette, and Marten Manimau 2011. *Dictionary of Kamang: Kamang – Indonesian – English*. Kupang, Indonesia: Kupang, Indonesia: Unit Bahasa dan Budaya (UBB)
- Steinhauer, Hein. 1995. Two varieties of the Blagar language (Alor, Indonesia). In Conny Baak, Mary Bakker, and Dick van der Meij, eds., *Tales from a concave world. Liber amicorum Bert Voorhoeve*: 265-296. Leiden: Department of languages and cultures of South-East Asia and Oceania, Leiden University.
- Steinhauer, Hein. Forthcoming. Blagar. In Antoinette Schapper, ed., *Papuan languages of Timor-Alor-Pantar: Sketch grammars Volume I*. Asia-Pacific Linguistics, Mouton de Gruyter.
- Stokhof, W.A.L. 1975. *Preliminary notes on the Alor and Pantar languages (East Indonesia)*. Canberra: Pacific Linguistics Series B-43.

6 Notes to the Appendices

Appendix I presents the AP data supporting the additional and revised pAP reconstructions; Appendix II gives the TK data supporting the additional and revised pTIM reconstructions; and finally, Appendix III gives a list of the 89 pAP and pTIM cognates, together with reconstructed pTAP forms, on which the sound correspondences identified in this chapter are based.

The AP data presented in Appendix I derive from the comparative lexical database compiled by the authors of Holton et al. (2012), as well as from Schapper’s field notes on Kamang and Wersing. Other than Holton et al. (2012), which bases its reconstruction on a subset of 12 AP languages, we give a full list of the AP languages found in the AP lexical database. This allows us to draw on cognate forms found only in some of the smaller languages. Note that for reasons of space, we abbreviate the language names; see below for an explanation of these abbreviations.

The TK data presented in Appendix II rely on Schapper’s and Huber’s field notes on Bunaq and Makalero, respectively. The Makasae data are drawn from Huber’s field notes as well as a number of online resources. The sources of both AP and TK language data are detailed below.

Sources

Abui (ABU)	Kratochvíl 2007, Kratochvíl and Delpada 2008, Schapper fieldnotes 2010	Kui (KUI)	Holton fieldnotes 2010
Adang (ADG)	Robinson fieldnotes 2010	Kula (KUL)	Holton fieldnotes 2010, Nicholas Williams p.c. 2011
Blagar (BLG)	Robinson fieldnotes 2010	Makalero	Huber 2011, Huber fieldnotes 2007-2013
Bunaq (Lamaknen)	Schapper n.d. a, Schapper 2010	Makasae	Brotherson 2003, Carr 2004, Huber 2008, Huber fieldnotes 2005, 2012-2013, Language Documentation Training Center of the University of Hawaii ⁴
Deing (DEI)	Robinson fieldnotes 2010	Nedebang (NED)	Robinson fieldnotes 2010
Fataluku	Fataluku online dictionary ⁵ , van Engelenhoven fieldnotes	Oirata	Josselin de Jong 1937, van Engelenhoven fieldnotes
Hamap (HMP)	Robinson fieldnotes 2010	Reta (RET)	Robinson fieldnotes 2010
Kamang (KMG)	Schapper n.d. b, Schapper and Manimau 2011	Sar (SAR)	Robinson fieldnotes 2010
Kabola (KAB)	Robinson fieldnotes 2010	Sawila (SWL)	Kratochvíl n.d.
Kaera (KAE)	Klamer Kaera corpus 2005-2007	Teiwa (TEW)	Klamer Teiwa corpus, Klamer and Sir 2011, Robinson fieldnotes 2010
Kafoa (KAF)	Baird fieldnotes 2003	Wersing (WER)	Schapper and Hendery fieldnotes 2012, Holton fieldnotes 2010
Klon (KLN)	Baird fieldnotes 2003	Western Pantar (WP)	Holton and Lamma Koly 2008, Holton fieldnotes 2010

The orthographic conventions used in the Appendices are the following: ‘~’ joins morphological variants of the same lexeme. In Appendix I and Appendix II, material given in round brackets ‘()’ represents fossilized morphology or other unetymological material. In Appendix III, round brackets indicate that a given phoneme cannot be reconstructed with absolute certainty. Furthermore, ‘N’ is used to represent an unspecified nasal; ‘L’ an unspecified liquid, and ‘Q’ a putative postvelar stop for which we have only very weak evidence. An empty slot in the pTAP column means that the reconstructed pAP and pTIM forms, although clearly cognate, are too different to allow for a secure pTAP reconstruction.

⁴ Online at http://www.ling.hawaii.edu/ldtc/languages/makasae_fatum/ and http://www.ling.hawaii.edu/ldtc/languages/makasae_osor/.

⁵ Online at www.fataluku.com

7 Appendix I: data supporting the additional pAP reconstructions

gloss	bark	bird	bite	bone	clew, circle ³	stone	coconut	crawl	di
pAP original	--	*dVl	--	--	--		*wat(a)	--	*r
pAP new	*lVu	*(a)dVl	*(ta)ki	*ser	*maita		*wata	*er	*r
SAR	--	dal	--	--	--		wat	--	m
DEI	--	dal	--	--	--		wat	--	m
TEW	--	dai	--	--	--		wat	--	m
NED	--	daya	--	--	--		wata	--	m
KAE	--	--	--	--	--		wat	--	m
WP	lau	--	--	--	--		hatua	--	--
BLG	olovi	--	(ga)ki	--	--		vet	--	(i)
RET	lu	--	ki(-ki)	--	--		vat	--	(a)
ADG	lowo?	--	--	--	--		fa?	--	m
HMP	--	--	--	--	--		--	--	m
KAB	olowo	--	--	--	--		wa?	--	m
KUI	--	adol	--	--	--		bat	--	m
KAF	--	--	--	--	--		--	--	(i)
KLN	--	--	--	--	--		--	--	--
ABU	lou	--	(ta)kai	--	masaŋ ɿ ⁴		wata	--	m
KMG	--	atul	ka(te) ¹	sɛl ɿ ²	maita		--	eei ~ eel	--
KUL	leloja	--	--	(gi)saja	--		g ^w ata	--	--
SWL	--	adala	--	sara	--		wata	--	--
WER	aloi	adol	(mi)kik	(ge)seri	--		wata	er	--

¹ Metathesised form; denotes ‘eat’. ² Kamang normally reflects pAP *r as as i in final position. ³ See Rodemeier (1992) on clews in Alor. ⁴ Abui normally reflects pAP *t as t. ⁵ This reconstruction must be viewed as tentative, since Kabola does not make part of the existing pAP reconstruction. ⁶ Note the loss of the initial syllable in several of the daughter languages. According to Holton et al. (2012) and Holton and Robinson (this volume), this has to do with stress being based on syllable weight. The heavy *bar syllable attracts stress, which leads to the loss of the initial syllable. A similar case is, possibly, pAP *tei ‘tree’.

gloss	ear	face	far	fire	fish	flat	girl	gran
pAP original	*-uar(i)	--	--	*had(a)	*hab(i)	--	--	--
pAP new	*-uari	*-pona	*lete	*hada	*habi	*tatok	*pon	*tar
SAR	--	--	--	--	--	--	--	--
DEI	-war	--	--	--	--	--	--	--
TEW	-uar	--	--	[127?]ar	[127?]af	--	--	--
NED	-ow	--	--	ar	a:fi	--	--	--

KAE	-uar	--	--	ad	ab	--	--	--
WP	-ue	--	--	--	hap	--	--	--
BLG	-veli	--	--	a:d	a:b	--	--	--
RET	--	--	--	--	--	--	--	--
ADG	--	--	--	--	a:b	--	--	--
HMP	--	--	--	--	--	--	--	--
KAB	--	--	--	--	--	--	--	--
KUI	-uel	--	--	ar	eb	--	--	--
KAF	--	--	--	--	--	--	--	--
KLN	-uer	--	--	əda	əbi	--	--	--
ABU	-uei	-poŋ	--	ara	afu	--	--	--
KMG	-uai	-funa:	letei	ati	api	tatok	fon	dum
KUL	--	--	--	--	--	--	--	atar
SWL	uari	--	--	ada	api	--	--	(ga)
WER	-ueri	--	--	ada	api	--	--	(ne)

¹ This is a reciprocal term. The reflexes in the modern languages denote either ‘grandparent’ or ‘grandchild’. ² Semantic shift to ‘child’. ³ Denotes ‘grandchild’. ⁴ Denotes ‘grandparent’. ⁵ While clearly cognate, the forms in this set show a variety of unexpected or irregular sound changes: Teiwa, Nedebang and Kaera normally reflect pAP *l as l in initial and medial position, rather than j; Teiwa and Nedebang normally reflect pAP *g as [127?] and x, respectively, in medial position, rather than g; pAP *g is normally reflected as g in Klon and j in Sawila; and finally, initial h in Western Pantar is usually a reflex of pAP *h, rather than *w. The pAP reconstruction must thus be seen as somewhat tentative.

gloss	itchy	laugh	leg	LOW	meat	mountain	name	ne
pAP origi- nal	--	*jari	--	--	--	--	*-ain(i, u)	*s
pAP new	*(i)ruk	*jagir ¹	*-bat	*po ²	*iser ³	*buku	*-en(i, u)	*s
SAR	--	jehar	-fat	--	--	--	--	--
DEI	--	jaxar	-wat	--	--	--	--	si
TEW	--	jə[127?]ar	-fat	--	--	--	--	si
NED	--	gela	--	--	--	--	-einu	sa
KAE	--	agar	at	--	--	buku:	-en	si
WP	--	jali ɿ	--	--	--	--	-inru	sa
BLG	--	iriga	--	po	--	buku	-ene	hi
RET	--	agala	--	--	--	--	--	ha
ADG	--	--	--	pə	hiri ɿ	--	-aniŋ	ha
HMP	--	--	--	--	(ma)hil	--	ane	ha
KAB	--	ja:la	--	--	--	--	--	--
KUI	rok	jeri ɿ	--	--	Is	--	-enei	sa
KAF	--	--	--	--	(ma)he:l	--	-nei	hi
KLN	--	əgar	--	--	(mə)hel	--	-əne?	ha

ABU	jokuŋ	--	--	pa	mahitiŋ	buku ɿ ⁴	-ane	tu
KMG	jokuŋ	--	--	fe	isei	buk ɿ ⁴	-nei	su
KUL	joka	geja	--	--	--	--	--	tu
SWL	--	jara ɿ	--	--	isi ɿ	--	-ani	tiŋ
WER	iruk	jer ɿ	--	--	(ge)is ɿ	--	--	tə

¹ Holton et al. (2012) reconstruct *jari for ‘laugh’. We revise this form on the basis of the clear presence of a medial velar in the reflexes of many AP languages. Note, however, the irregular loss of reflexes of pAP *g in Western Pantar, Kui, Sawila and Wersing. ² See Schapper (this volume) for details on this reconstruction. ³ The reflexes of this form denote ‘game’ or ‘meat’. Note that there are several irregularities in this set: Adang normally reflects pAP *r as l, rather than r; and Sawila and Wersing normally reflect *s as t, rather than s. ⁴ Abui normally reflects pAP *b as f, rather than b, and pAP *b is usually reflected in Kamang as p, rather than b. ⁵ Denotes ‘coast’. The relationship between the two senses is explained by the typical settlement patterns in the region: older settlements are located in high places, often on top of knolls or ridges, whilst newer settlements are downhill towards the coast.

gloss	path	person	price	<i>P. indicus</i> ⁴	rain	ripe	scratch	sh
pAP original	--	--	--	--	--	*ten	--	--
pAP new	*jega ¹	*anin	*bol ²	*matar	*anur	*tena	*karab	*s
SAR	--	--	--	--	--	--	kəra:b	si
DEI	--	--	--	--	--	ten:raŋ	krab	si
TEW	--	--	--	--	--	--	--	si
NED	ji:ja ɿ	--	--	--	--	tanana	(ki)kar ɿ ⁵	--
KAE	--	--	--	--	--	ten-	krabis ɿ ⁶	si
WP	ja ɿ	--	--	mat:ɛ	--	taŋ	karasi ɿ ⁷	si
BLG	iga ɿ	--	--	--	onor	tena	--	si
RET	viag	--	(ta)beli ³	--	--	--	--	hi
ADG	se?	--	--	--	nui	tene	--	--
HMP	se?	--	--	--	--	tən	--	--
KAB	je?	--	(ʔo)wol ³	--	nui	tenaŋ	--	--
KUI	--	anin(ou)	--	--	anor	tain	ukuberi	so
KAF	ʔije	--	--	--	--	--	ukafi	--
KLN	ɛge?	anin(ok)	--	mtar	--	ətən	kərɔb	--
ABU	--	--	(he)bel ³	mitai	anui	--	kafi	--
KMG	--	--	bol ³	--	--	iten ~ iton	--	--
KUL	--	aniŋ(na)	--	--	--	--	kapi	--
SWL	--	aniŋ(ka:)	--	mata:ri	--	itima	kapari	--
WER	--	aniŋ	--	--	--	--	kəpir	--

¹ There are a number of irregularities in this set: Nedebang normally reflects medial *g as x, Western Pantar as g:, and Blagar as either Ø or [241?]. ² This root is likely an Austronesian loan: PMP *bəli ‘price’, bride price’. ³ Denotes ‘bride price’. ⁴ New Guinea rosewood (*Petrocapus indicus*), typically referred to in eastern Malay as *kayu merah*. ⁵ Note the irregular loss of the final syllable. ⁶ Semantic shift to ‘claw’. Also, note the unetymological s, present in both Kaera and Western Pantar. ⁷ While this form is very likely related, it includes several irregularities: the expected reflex of pAP *r in medial position is l in Western Pantar; there is no reflex of pAP *b, which is normally reflected as b; and there

is an unetymological s.⁸ Blagar normally reflects pAP *s as h in word-initial position.⁹ This set shows a variety of irregularities: Adang normally reflects pAP *r as l or I, rather than r; pAP *r is normally reflected as i in final position in both Abui and Kamang; and Wersing normally reflects pAP *s as t, rather than s.

gloss	stand	sugarcane	sun	taboo	tail	tongue	tooth	w
pAP origi- nal	--	*u:b	*wad(i)	--	*-or	*-leb(ur)	*-uas	--
pAP new	*nate(r) ¹	*hu:ba	*wadi	*palol	*-ora	*-lebur	*-uasin	*l
SAR	--	--	war	--	-or	--	--	--
DEI	--	--	--	--	-or	--	--	--
TEW	--	--	war (get)	--	-or	-livi	-usan	--
NED	--	u:fa	weri	--	-ola	-lefu	-usiŋ	--
KAE	--	u:b	wer	--	-or	-le:b	-uasŋ	--
WP	natar ɿ ²	--	war	--	--	-lebu	-wasŋ	--
BLG	--	ub	ved	--	ora	-dʒebur	-veŋ	--
RET	--	juwab	vid	--		-lebul	--	--
ADG	--	so:b	fɛd	--	oloʔ	-lɛb	-wɛheŋ	--
HMP	--	--	fɔd	--	ol	--	-fiʔiŋ	--
KAB	--	job	wer	--	ʔol	-leb	--	la
KUI	--	u:b	ber	--	-or	-liber	-wes	la
KAF	natei	--	uru	--	--	-lip	-weheŋ	la
KLN	--	--	--	--	-or	-lɛb	-wɛh	--
ABU	nate	fa	wari	palol	--	-lifi	-weiti	la
KMG	--	--	wati	foi	-(w)ui	-opei	-weh	lo
KUL	--	p ^w a	wad	--	--	ilip	--	--
SWL	--	--	wadi	--	-(w)o:ra	--	-wa	--
WER	--	upa	widi	--	wori	-jebur	-wesi	--

¹ There is a competing and morphologically unrelated form *tas ‘stand’, which is more widely distributed across modern AP languages (see Holton and Robinson, this volume). ² Western Pantar normally reflects pAP *r as Ø in word-final position. ³ This root is possibly an Austronesian loan: PMP *lakaj ‘stride, take a step’. ⁴ Kamang normally reflects pAP *k as k. ⁵ Semantic shift to ‘follow’. ⁶ Kaera normally reflects pAP *l as l in word-initial position.

gloss	yellow
pAP origi- nal	--
pAP new	*bagori
SAR	bahar
DEI	bug
TEW	ba[127?]ari
NED	baxori
KAE	bagari

WP	bugra
BLG	bagori ɟ ¹
RET	bagori
ADG	baʔoi
HMP	baʔoil
KAB	baʔoil
KUI	bagura
KAF	fijoi
KLN	bubugər
ABU	--
KMG	--
KUL	--
SWL	--
WER	--

¹ Blagar normally reflects pAP *g as Ø or [241?] in medial position.

8 Appendix II: data supporting the additional pTIM reconstructions

gloss	banana	bark	bat	bite	bone	breast	child
pTIM original	*muku	--	--	*gakel	--	--	--
pTIM new	*mugu	*le(k)u(l)	*maTa	*(ga)gel	*(se)sa(r, R)	*hami	*-al
BUNAQ	mok	--	--	gagil	sesal	--	-ol
MAKASAE	muʔu	leu ¹	--	gaʔel	--	ami	--
MAKALERO	muʔu	leu ¹	--	kaʔel	--	--	--
FATALUKU	muʔu	leʔul(e) ²	maca	(ki)kiʔ(e)	--	ami(-tapunu) ₃	--
OIRATA	mu:	leule ²	maʔa	--	--	--	--

¹ Semantic shift to ‘call’. ² Semantic shift to ‘sing’. ³ This lexeme is a lexical doublet, i.e. originally a compound or a lexicalized parallel expression (see Schapper et al. 2012: 224). ⁴ Semantic shift to ‘bent over (as with age)’. ⁵ This form shows metathesis in Proto-Eastern Timor: *kari > *raki > raʔi / raʔe(ne). ⁶ Semantic shift to ‘littered with stones’.

gloss	dream	eat	excrement	face	far	fish	flat	ga
pTIM original	--	--	--	*fenu	--	*api	--	*(
pTIM new	*ufar(ana)	*nua	*a(t, D)u	*-fanu	*eTar	*hapi	*tetok	*(
Bunaq	waen ¹	a ~ -ia	ozo	-ewen	ate	--	toiʔ ⁴	m
Makasae	ufarena	nawa	atu[-guʔu] ²	fanu	--	afi	--	ar
Makalero	ofarana	nua	atu	fanu	--	afi	tetuʔ	ar
Fataluku	ufarana	una ~ naβa	atu ³	fanu	icar	api	--	--
Oirata	upar(a)	una ~ nawa	atu ³	panu	--	ahi	--	un

1 This item shows metathesis: waen < *awen following on fusion from the two halves of the reconstructed doublet. 2 The Bunaq cognate for the second half of this lexical doublet is *g-io* ‘3AN-faeces’, but it doesn’t appear in a doublet with *ozo* ‘faeces’. 3 Semantic shift to ‘belly’. 4 The final glottal stop in Bunaq is likely a reflex of final *k in pTIM. However, more evidence is needed to substantiate this claim.

gloss	laugh	leg	LOW	mat	mountain	new	new place	no
pTIM original	*hika	--	--	--	--	*(t, s)ifa	--	--
pTIM new	*jiger	*-buta	*ufe	*biti	*bugu	*(t, s)ipa(r)	*lan	*-
Bunaq	higal	-but ¹	--	--	--	tip	lon	-in
Makasae	hiʔa	--	he- ʔ ²	--	buʔu	sufa	--	m
Makalero	hiʔe	--	ufe-	piti	puʔu ³	hofar	--	m
Fataluku	heʔe	--	ua- ʔ ²	pet(u)	--	--	--	m
Oirata	--	--	ua ʔ ²	het(e)	--	--	--	--

¹ Semantic shift to mean ‘knee’. ² The reflex of pTIM *f as /h/ in Makasae and Ø in Fataluku and Oirata is irregular; /f/ is expected for Makasae and Fataluku, and /p/ for Oirata. ³ Semantic shift to ‘gable, top of house’. ⁴ This item appears to show metathesis in the following stages: pTIM *-muni > *-minu > *-imun > *-inum > Bunaq *-inup* ‘nose’. The change of *m to Bunaq p is explainable as the

result of *m* being prohibited from codas in Bunaq. ⁵ The suffix *-kai* is frequently found in body part terms in Makasae. ⁶ It seems likely that medial **p* changes to /w/ in Bunaq. However, we currently lack sufficient data to support this conclusion. There has also been a semantic shift to ‘foreigner’.

gloss	path	person	rain	ripe	scorpion	scratch	shark	si
pTIM original	*hika	--	--	*tina(k)	--	--	--	--
pTIM new	*jiga	*anu	*ine(r, R)	*tena	*fe(r, R)e	*gabar	*supor	*t
Bunaq	hik	en	inel	ten ¹	wele ⁴	--	--	to
Makasae	hiʔa	anu	--	tina ²	--	--	--	--
Makalero	hiʔa	anu	--	tina ~ dina ²	--	kapar	su(-amulafu) ₅	--
Fataluku	iʔa	--	--	tina ³	--	kafur(e)	hopor(u) ⁶	--
Oirata	ia(ra)	--	--	--	--	--	--	--

¹ Semantic shift to ‘be cooked, ready’. ² Semantic shift to ‘cook’. ³ Semantic shift to ‘set alight’. ⁴ It seems likely that initially before front vowels **f* changes to /w/ in Bunaq. However, we currently lack sufficient data to support this conclusion. ⁵ The meaning of the compound *su-amulafu* is not quite clear. It seems to refer to a large sea creature, possibly a dolphin or a dugong. The second element, *amulafu*, translates as ‘human being, person’. ⁶ This form is glossed in various ways in the different Fataluku sources either as ‘shark’ or ‘dugong’

gloss	tree	walk 1	walk 2	yellow	1SG	1PI
pTIM original	*hote	*lakor	--	--	--	--
pTIM new	*hate	*lagar ¹	*male	*gabar	*n-	*fi
Bunaq	hotel	lagor	mele	--	n-	--
Makasae	ate	laʔa	--	gabar	--	fi
Makalero	ate	laʔa	--	--	--	fi
Fataluku	ete	laʔa	--	--	--	afi
Oirata	ete	lare	--	--	--	ap-

¹ This root is possibly an Austronesian loan: PMP **lakaj* ‘stride, take a step’.

9 Appendix III: list of cognates and pTAP reconstruction

gloss	pTAP	pAP	pTIM
bamboo	*mari	*mari	*mari
banana	*mugul	*mogol	*mugu
bark, call		*lVu	*le(k)u(l)
bat	*madel	*madel	*maTa
bathe	*weLi	*weli	*weru
bird	*(h)adul	*(a)dVl	*haDa
bite	*ki(l)	*(ta)ki	*(ga)gel
blood	*waj	*wai	*waj
bone	*se(r, R)	*ser	*(se)sa(r, R)
breast	*hami	*hami	*hami
grandparent	*(t, d)ama	*tam(a, u)	*moTo
child	*-uaQal	*-uaqal	*-al
clew	*ma(i)ta(r)	*maita	*matar
coconut	*wata	*wata	*wa(t, D)a
crawl	*er	*er	*er
crouch	*luk(V)	*luk(V)	*luk
die	*mV(n)	*min(a)	*-umV
dirty	*karV(k)	*karok	*gari
dog	*dibar	*jibar	*Depar
dream	*(h)ipar	*hipar	*ufar(ana)
ear	*-waRi	*-uari	*-wali
eat	*nVa	*nai	*nua
excrement	*(h)at(V)	*has	*a(t, D)u
face	*panu	*-pona	*-fanu
far	*le(t, d)e	*lete	*eTar
fire	*hada	*hada	*haTa
fish	*habi	*habi	*hapi
flat	*tatok	*tatok	*tetok
garden	*magad	*magad(a)	*(u, a)mar
girl	*pan(a)	*pon	*fana
give	*-(e, i)na	*-ena	*-inV
green	*lugar	*(wa)logar	*ugar
hand	*-tan(a)	*-tan	*-tana
hear	*ma(g, k)e(n)	*magi	*mage(n)
inside	*mi	*mi	*mi
itchy	*iRak	*(i)ruk	*ilag

laugh	*jagir	*jagir	*jiger
leg	*buta	*-bat	*-buta
LOW	*po	*po	*ufe
mat	*bi(s, t)	*bis	*biti
meat	*isor	*iser	*seor
moon	*hur(u)	*wur	*huru
mountain	*buku	*buku	*bugu
name		*-en(i, u)	*-nej
new	*(t, s)iba(r)	*siba(r)	*(t, s)ipa(r)
new place	*lan	*lan	*lan
nose	*-mVN	*-mim	*-muni
one	*nukV	*nuk	*uneki
other	*abe(nVC)	*aben(VC)	*epi
<i>P. indicus</i>	*matar	*matar	*ma(t, D)ar
path	*jega	*jega	*jiga
person	*anV(N)	*anin	*anu
pig	*baj	*baj	*baj
pound	*tapa(i)	*tapai	*tafa
price	*boL	*bol	*bura
rain	*anu(r, R)	*anur	*ine(r, R)
rat	*dur(a)	*dur	*Dura
ripe	*tena	*tena	*tena
run	*tipar	*tiara	*tifar
scorpion	*pV(r, R)	*pVr	*fe(r, R)e
scratch	*karab	*karab	*gabar
sea	*tam(a)	*tam	*mata
shark	*sibar	*sib(a ,i)r	*supor
sit	*mit	*mis	*mit
six	*talam	*talam	*tamal
sleep	*tia(r)	*tia	*tia(r)
spit	*puRV(n)	*purVn	*fulu(k, n)
spoon	*suRa	*surV	*sula
stand	*nat(er)	*nate(r)	*nat
star	*jibV	*jibV	*ipi(-bere)
stone	*war	*war	*war
sugarcane	*ub(a)	*hu:ba	*upa
sun	*wad(i, u)	*wadi	*waTu
taboo	*palu(l, n)	*palol	*falu(n)
tail	*-oRa	*-ora	*-ula(?)

tongue	*-lebuR	*-lebur	*-ipul
tooth	*-wasin	*-uasin	*-wasin
tree	*hate	*tei	*hate
vagina	*-ar(u)	*-ar	*-aru
wake	*tan(i)	*-ten	*Tani
walk 1	*lak(Vr)	*laka	*lagar
walk 2	*lamV	*lam(ar)	*male
water	*jira	*jira	*ira
weave	*sine(N)	*sine(N)	*sina
yellow	*bagur(V)	*bagori	*gabar
1PI	*pi	*pi-	*fi
1SG	*na-	*na-	*n-
3ALN	*gie	*ge	*gie
3INLN	*g(a, i)-	*ga-	*g-
