The Historical Relations of the Papuan Languages of Alor and Pantar

Gary Holton,* Marian Klamer,† František Kratochvíl,‡
Laura C. Robinson,* and Antoinette Schapper†

*UNIVERSITY OF ALASKA FAIRBANKS, †LEIDEN UNIVERSITY,
AND ‡NANYANG TECHNOLOGICAL UNIVERSITY

The historical relations of the Papuan languages scattered across the islands of the Alor archipelago, Timor, and Kisar in southeast Indonesia have remained largely conjectural. This paper makes a first step toward demonstrating that the languages of Alor and Pantar form a single genealogical group. Applying the comparative method to primary lexical data from twelve languages sampled across the islands of the Alor-Pantar archipelago, we use form-meaning pairings in basic cognate sets to establish regular sound correspondences that support the view that these languages are genetically related. We reconstruct 97 Proto–Alor-Pantar vocabulary items and propose an internal subgrouping based on shared innovations. Finally, we compare Alor-Pantar with Papuan languages of Timor and with Trans-New Guinea languages, concluding that there is no lexical evidence supporting the inclusion of Alor-Pantar languages in the Trans-New Guinea family.

1. INTRODUCTION.[1] The historical relations of the Papuan languages scattered across the islands of the Alor archipelago, Timor, and Kisar in southeast Indonesia have remained largely conjectural.[2] This paper takes a step toward an empirical demonstration of the mutual relatedness of the languages of Alor, Pantar, and the intervening islands of Tereweng, Pura, and Ternate (see figure 1). Applying the comparative method to primary lexical data from twelve languages sampled across the islands, we use form-meaning pairings in 109 basic cognate sets (78 of which we reconstruct to the level of Proto–Alor-Pantar) to establish regular sound correspondences that attest to the genealogical relationship between these languages.

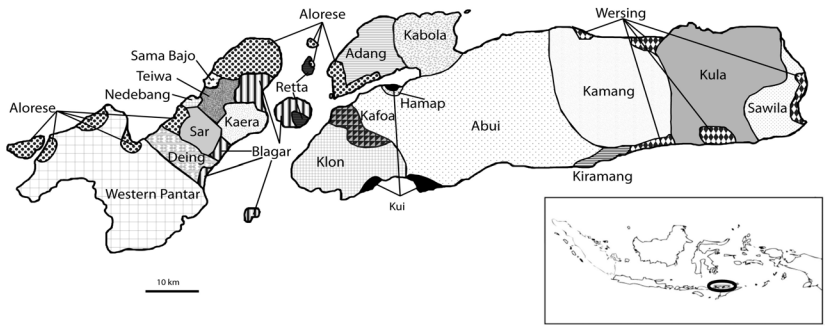
1. Field work was supported by grants from the Netherlands Organization for Scientific Research, the UK Arts and Humanities Council, and the US National Science Foundation (NSF-SBE 0936887), under the aegis of the European Science Foundation EuroBABEL program and Nanyang Technological University, Singapore. Authors' names appear in alphabetical order. We thank two anonymous reviewers for many insightful comments; any remaining errors of fact or interpretation are our own.

2. The term "Papuan" is generally used as a cover for the perhaps 800 languages spoken in New Guinea and its vicinity that are not Austronesian (cf. Ross 2005:15). It says nothing about the genealogical ties between languages. In this paper we use "Papuan" as synonymous with "non-Austronesian" to refer to any language that is spoken in the area of New Guinea (excluding Australia) but is not a member of the Austronesian language family.

The first published body of lexical information on the Papuan languages of the islands of the Alor archipelago (henceforth "AP languages") was Stokhof (1975). This work contains 117-item basic word lists for 34 villages representing twelve AP language varieties. It was followed by the appearance of several works on individual AP languages of variable depth and scope: Stokhof (1977, 1979, 1982) on Kamang (Woisika), Stokhof (1984) on Abui, Steinhauer (1991, 1995) on Blagar, and Donohue (1996) on Kula (Tanglapui). The amount of lexical data on the AP languages, however, remained scant. From 2003 onward, there has been a surge in descriptive work and more detailed surveying of the AP languages. Recent and forthcoming publications on AP languages include Baird (2005, 2008, 2010), Holton (2004, 2008, 2010), Holton and Lamma Koly (2008), Klamer (2008, 2010a–c, 2011, 2012, forthcoming:a,b), Klamer and Kratochvíl (2006, 2010), Klamer and Schapper (forthcoming), Kratochvíl (2007), Kratochvíl and Delpada (2008), Schapper and Manimau (2011), and Schapper and Klamer (2011). This, as well as extensive additional detailed surveying that took place in 2010, provides the data for our investigation of the historical relations between the AP languages (see also section 3). This paper reports the first step in that investigation.

In the literature, the Alor-Pantar languages have usually been discussed together with the Papuan languages of Timor and Kisar, to which they are surely related (see discussion in the following section). However, in this paper we restrict our attention to the reconstruction of Proto–Alor-Pantar (PAP), for two primary reasons. First, our data set for the AP languages is most robust, being based on several years of first-hand field work by the authors and colleagues. Second, while there remains much disagreement as to the nature of genetic affiliations between the Timor-Kisar and the Alor-Pantar languages, it is clear that the latter form a well-defined group. Lexical correspondences are abundant and more or less regular, and there is much shared grammatical structure. In contrast, the Timor-Kisar languages reflect significant lexical and grammatical influence from neighboring Austronesian languages. This situation largely reflects a distinct sociolinguistic history in Alor-Pantar. Until the modern era, only a single Austronesian language, Alorese, managed

FIGURE 1. MAP OF THE ALOR-PANTAR LANGUAGES



to get a foothold in the Alor-Pantar archipelago, and Austronesian speakers never settled in the interior regions.[3] Given the integrity of AP as a group, a reconstruction of PAP is a necessary prerequisite to assessing wider genetic affiliations.

This paper is structured as follows. Section 2 gives a brief background on the classification of Papuan languages in East Nusantara, including the AP languages. Section 3 presents details of the data used in this paper. Section 4 discusses the sound correspondences and preliminary phoneme reconstructions. We limit ourselves here to the reconstruction of consonant phonemes. Section 5 looks at what sound changes are shared by contemporary AP languages, and the subgrouping evidence that this affords. Section 6 briefly explores the implications of these findings and questions for further investigation.

2. HISTORY OF CLASSIFICATION OF THE PAPUAN LANGUAGES OF SOUTHEAST NUSANTARA. The non-Austronesian character of the languages of the Alor-Pantar archipelago (defined as the islands of Alor and Pantar together with the small islands in the intervening Pantar Strait) and that of several languages on neighboring Timor and Kisar has long been recognized in the literature. Early reports focused on racial and cultural distinctions. By the early twentieth century, Dutch military and scientific expeditions had already noted a clear distinction between the "nonindigenous" coastal populations (today's speakers of the Austronesian language Alorese) and the "indigenous" mountain populations of Alor and Pantar (Anonymous 1914:75–78).[4] Similarly, Vatter (1932: 278–79) observed a clear distinction between the people, cultures, and languages of East Flores and Solor on the one hand, and those of Alor and Pantar on the other, though he assumed that the languages belonged to the same "Indonesian" group. The earliest reference to the linguistic distinctiveness of this region is found in de Josselin de Jong's (1937) description of the Oirata language on Kisar (east of Timor). De Josselin de Jong describes Oirata as a close relative of Fataluku, spoken on the eastern tip of Timor. This insight led Nicolspeyer (1940) to suggest in her dissertation on Abui social structure that this central Alor language was also non-Austronesian.[5] Shortly thereafter, Capell (1943b) identified Bunaq and Makasae on Timor as Papuan.[6]

Very little new information about the Papuan languages of this region appeared in the literature in the following three decades. Capell (1975:673) remarked on the difficulty of characterizing these languages as a genealogical grouping, given that "Abui is only one language of Alor and the rest still remain unchronicled: they may be similar or they may not." The first clear indication that the Papuan languages of Alor-Pantar form a genealogical group is found in Stokhof's (1975) survey of basic vocabulary. Stokhof compares short (117-item) word

3. Alorese (Bahasa Alor, ISO 639-3 aol, Klamer 2011), an Austronesian language spoken along the northwestern coasts of Alor and Pantar, is a close relative of Lamaholot to the west of Pantar.
4. A footnote in this article explains that its two major sources were (i) the "Military Memories" that report on Dutch military expeditions on the islands in 1910 and 1911, and (ii) a report of a geological expedition by R. D. M. Verbeek in 1899, published as "Molukken Verslag" in the *Jaarboek van het Mijnwezen in Ned. Oost-Indie*, 1908.
5. Nicolspeyer's information on Abui is based on the extensive fieldwork conducted by Cora Du Bois in 1938 and 1939 in Abui-speaking Atimelang. Du Bois corresponded with de Josselin de Jong on Abui, and the latter noted that Abui did not seem to have an Austronesian lexicon.
6. Capell actually listed two additional languages as Papuan: Waimaha and Kairuhi; both have since been shown to be Austronesian.

lists from 34 villages in Alor-Pantar, concluding that, with the exception of the Austronesian language Alorese, "the Pantar-Alor languages are lexicostatistically related" (1975:22). Using a 70 percent threshold for percentage of shared cognates, Stokhof classifies these 34 varieties into 12 distinct Papuan languages, plus Austronesian Alorese.[7] No further discussion of the internal relationships of the Alor-Pantar languages is found in the literature. Instead, most of the ensuing comparative work on the Papuan languages of southeast Nusantara has simply assumed their relatedness, in order to focus on higher order classifications.

The wider genetic affiliations of the Alor-Pantar languages have been the subject of much speculation and hypothesis. The connection with the Papuan languages of Timor and Kisar—namely Bunaq (ISO 639-3 bfn), Makasae (mkz), Fataluku (ddg), and Oirata (oia)—has been widely assumed, at least in part due to their geographic proximity to Alor-Pantar. Stokhof (1975) also cites their typological affinity. Further, based on lexicostatistical evidence, Stokhof argues for a closer affinity between Makasae and the Alor-Pantar group (1975:24). In contrast, Ross groups Bunaq with Alor-Pantar, opposed to an East Timor group, noting that the latter group shares an innovation in the first person singular pronoun, namely, *ani* < Proto–Trans-New Guinea (PTNG) *na (Ross 2005:36). More recent morphological evidence from Timor languages suggests that the Papuan languages of the Timor group including Bunaq may be more closely related to one another than to the Alor-Pantar languages (Schapper 2010:21, 346).

The first proposals for wider genetic affiliations beyond the geographically immediate region compared Timor-Alor-Pantar (TAP: that is, AP plus the Timor-Kisar languages) to the languages of North Halmahera, the only other Papuan outlier languages spoken in the eastern Indonesian archipelago at a distance from the New Guinea mainland. Capell (1944) noted similarities between the Papuan languages of Timor and those of North Halmahera but initially refrained from asserting a genetic relationship. Anceaux (1973), commenting on a fieldwork report from the Pantar language Teiwa (Watuseke 1973), proposed including Teiwa and several Alor languages (Abui, Wersing, Kui) with Cowan's (1957) West Papuan group.[8] As later formalized, Capell's (1975) West Papuan Phylum included the "Alor-Timor" languages. In fact, only one Alor language, Abui, was included in Capell's grouping, as Capell only belatedly became aware of Watuseke's Teiwa data and the other extant Alor sources mentioned by Anceaux. Even with these additional data, Capell was well aware of the tenuous nature of the relationship between TAP (actually Alor-Timor) and North Halmahera, particularly the lack of identifiable lexical correspondences. He thus proposed a major split between Alor-Timor on the one hand, and the rest of the West Papuan Phylum on the other. Stokhof suggested connecting TAP with several languages of the Western Bird's Head of New Guinea, concluding that "the Alor-Pantar languages form a closely related group with Cowan's West Papuan Phylum" (1975:26). However, the puta-

7. Stokhof's survey overlooks a small community of Bajau (ISO 639-3 bdl) speakers on Pantar that probably arrived before the beginning of the nineteenth century, based on their own reports and confirmed as likely by James Fox (pers. comm., 2010.). Speakers on Pantar report that additional Bajau speakers immigrated to Pantar in the 1950s. A more recent group arrived from East Timor in 1999 (Klamer 2010:13).

8. Curiously, Anceaux refers to Cowan's West Papuan group as "demonstrated," while Cowan himself is much more cautious, noting merely that "the relationship with NH [i.e., North Halmahera] can at least be said to be a real possibility" (1957:91).

tive West Papuan languages with which Stokhof compared Alor-Pantar were later reclassified as Trans-New Guinea, rendering this lexical evidence moot.

As new data from Alor-Pantar and from languages in western New Guinea became available, the proposed connection between TAP and the West Papuan Phylum began to unravel. Capell's (1975) paper on the West Papuan Phylum was published with an editorial preface noting that the TAP languages should instead be included within the Trans New Guinea Phylum (Wurm, Voorhoeve, and McElhanon 1975). However, Wurm, Voorhoeve, and McElhanon (1975) provided no data to back up this classification, and instead remained skeptical as to whether TAP should be classified as Trans-New Guinea or West Papuan. In particular, they asserted that "whichever way they [the Timor-Alor-Pantar languages] are classified, they contain strong substratum elements of the other … phyla involved" (1975:318). In spite of this hedge, the classification of TAP languages as Trans-New Guinea continues to be assumed by many authors.

With the revival in interest in comparative Papuan linguistics at the beginning of the twenty-first century (cf. Pawley 2005), some attempts were made to connect the TAP languages to the Trans-New Guinea Phylum using the comparative method. Pawley (2001, n.d.) provides 98 PTNG reconstructions; of these, nine are given with reflexes in TAP languages. Only one Alor-Pantar language, Blagar, is represented, and five of Pawley's PTNG reconstructions have Blagar reflexes. No attempt is made to posit sound changes that gave rise to these Blagar forms, and, in fact, the correspondences seem to be based more on subjective similarity than on regular sound change. For example, Pawley cites PTNG *(mb,m)elak 'lightning' > Blagar *merax*, but *pululu 'fly, flutter' > Blagar *alili*. In the latter, original intervocalic *l is preserved as *l*, whereas in the former, *l is reflected as *r.* It should be acknowledged that Pawley's primary goal was not to evaluate the position of TAP but rather to inspire a revival in Papuan comparative linguistics. The few TAP comparanda cited by Pawley do not support inclusion of TAP in the Trans-New Guinea Phylum. In section 6 we will revisit the lexical evidence for a connection between TAP and TNG in light of our proposed reconstruction.

Other attempts to classify TAP languages have focused on typological criteria. Ross (2005) groups TAP as part of his "West Trans New Guinea linkage," along with the West Bomberai, Wissel Lakes, and Dani families. Ross bases his classification on pronominal forms, specifically independent pronouns. However, while the match between first person pronouns in TAP and TNG is indeed quite good (PAP *na-, PTNG *na), the match in the second and third persons can only be derived via a flip-flop in which second person pronouns trade places with third person. Compare PTNG *ŋga 2PRO and *(y)a 3PRO with Teiwa *ha ʔan* and *ga ʔan*, respectively. As additional evidence, Ross cites the shared innovation of a first person plural pronoun *bi in both West Bomberai and TAP. However, as Ross clarifies in a footnote, the reflex of *bi is an exclusive pronoun in West Bomberai but an inclusive pronoun in TAP.

More recently, Donohue (2008) has revived the possibility of a connection between West Papuan and TAP, suggesting a history of contact between a West Papuan source language and a pre-TAP language spoken in the Bomberai peninsula and South Bird's Head area. Donohue notes the similarity between the Timor-Alor-Pantar distributive plural (12MINIMAL in Donohue's terminology) *ta and the Proto-North Halmaheran (PNH) first

person singular pronoun *to-. Additionally, Donohue cites two lexical similarities—Proto–Timor-Alor-Pantar (PTAP) *aDa, PNH *gota 'tree'; and PTAP *yar, PNH *gala 'water'—each with similar forms in the Bird's Head and the Bomberai Peninsula of mainland New Guinea. Donohue's proposal raises the possibility that the linguistic prehistory of the Alor-Pantar languages may evade a simple cladistic classification; however, given the limited evidence in this proposal for a West Papuan connection, it should be approached with some skepticism.

The current paper differs from previous research in several important ways. First, rather than simply comparing phonetically similar forms, we employ the comparative method. We focus on a close comparison of lexical forms to establish regular sound correspondences and to reconstruct protophonemes and the sound changes that they have undergone. Second, we restrict our attention to the Alor and Pantar languages, choosing to leave the exploration of possible more distant genetic affiliation with the Papuan languages of Timor for future investigation. Finally, the data set used in this paper is larger and more detailed than any of the data sets used in previous research, as discussed in the following section.

3. THE DATA. There are over twenty Papuan languages spoken across Alor, Pantar, and the islands in the intervening straits.[9] In this paper we present data from twelve languages (listed in), with representatives from across the entire geographical range of the area. The table includes the three-letter language abbreviation used in this paper, the name used by Stokhof (1975), the ISO 639-3 code, the name of the researcher who collected the data, the year(s) when collection took place, the number of lexical items in our data set, and an indication of the type of source.

Most of our data have been collected recently by the authors and colleagues during extensive periods of fieldwork. Some data come from sources published earlier. In some cases, the data present uncertainties regarding the phonemic status of particular segments, orthographic conventions, and morpheme boundaries. For example, in elicited word lists, verbs can occur with as yet unanalyzed aspectual and/or modal suffixes. In this paper, we only compare root forms, with affixes being identified on the basis of grammatical descriptions and recurrent endings within the lexical data. In the cognate sets presented here, material identified as fused or fossilized morphology is bracketed with ( ), while roots that obligatorily occur with affixes are marked with a hyphen.

Drawing from a comparative lexical database consisting of approximately 400 items, we identify 108 cognate sets reflecting regular sound correspondences. The range of semantic domains covered by the cognate sets and the number of sets represented in each domain are given in table 2. There are only 106 distinct meanings, as two of the meanings, 'dog' and 'walk', are found in more than one cognate set.

These forms show predominantly regular sound correspondences, as described in the following section. However, it is important to note that several of the cognate sets cannot be reconstructed as words in the protolanguage, since they are found only in a geographically restricted area (see section 5 below). That is, in some cases lexemes appear to have been innovated. This is particularly obvious for those meanings for which we have two corre-

9. The precise number of languages remains unknown, owing primarily to a lack of data for varieties spoken in the highlands of central Alor, which are currently classified as Abui or Kamang.

TABLE 1. SOURCES CONSULTED FOR THIS PAPER

LANGUAGE	ABBR.	STOKHOF (1975)	ISO CODE	RESEARCHER	YEAR(S)	NO. ITEMS	SOURCE(S)
Teiwa	TWA	Tewa	twe	Klamer	2003–7, 2010	1350	Klamer 2010a, forthcoming:a
				Robinson	2010	~400	fieldnotes
Nedebang	NDB	Nedebang	nec	Robinson	2010	~400	fieldnotes
Kaera	KER	—	—	Klamer	2006–7	890	fieldnotes, Klamer 2010b
				Robinson	2010	~400	fieldnotes
Western Pantar*	WP	Lamma	lev	Holton	2006–8	2500	Holton and Lamma Koly 2008
Blagar†	BLG	Blagar	beu	Robinson	2010	~400	fieldnotes
Adang	ADN	Adang	adn	Robinson	2010	~400	fieldnotes
				Baird	2003	419	fieldnotes
Klon	KLN	Kelon	kyo	Baird	2003–7	~1600	Baird 2008, fieldnotes
				Robinson	2011	~400	fieldnotes
Kui	—	Kui	kvd	Holton	2010	432	fieldnotes
Abui	ABU	Abui	abz	Schapper	2010	~400	fieldnotes
				Kratochvíl	2003–9	1725	Kratochvíl 2007, Kratochvíl and Delpada 2008
Kamang	KMN	Woisika	woi	Schapper	2010	~1800	fieldnotes
Sawila	SWL	Tanglapui	tpg	Kratochvíl	2007–9	~1800	fieldnotes
Wersing	WRS	Kolana	kvw	Holton	2010	432	fieldnotes

* Western Pantar is a cover label first used by Holton (2004) for three mutually intelligible dialects: Mauta, Tubbe, and Lamma (labels are based on clan names). The name "Lamma" is used in Stokhof (1975) for all varieties of the language.

† Blagar exhibits significant dialect variation with respect to the consonants. Unless otherwise noted, the data cited in this paper are from the Nuhawala "Nule" dialect spoken on Pantar.

TABLE 2. ALOR-PANTAR COGNATE SETS BY SEMANTIC DOMAIN

	TYPE	MEANINGS	NO.
A	**Pronouns**	1SG, 1PL.INCL, 2SG, 3SG, 3GEN, 3LOC, 3PL	7
	Nouns		
B1	Humans	'child', 'grandchild', 'name'	3
B2	Body Parts	'hand/arm', 'blood', 'body hair', 'bone', 'breast', 'ear', 'fat', 'fingernail', 'horn', 'leg', 'mouth', 'tail', 'teeth', 'tongue', 'wound'	15
B3	Environment	'fire', 'lime', 'moon', 'saltwater', 'sky', 'smoke', 'star', 'stone', 'sun' , 'water', 'wave', 'wind'	12
B4	Flora	'banana', 'betel nut', 'betel vine', 'coconut', 'maize', 'rattan', 'tree'	7
B5	Fauna	'bat', 'bedbug', 'bird', 'crocodile', 'dog', 'fish', 'flea', 'lizard', 'monitor lizard (*Varanus sp.*)', 'mosquito', 'pig', 'rat', 'scorpion', 'shark'	14
B6	Property / Culture	'axe', 'comb', 'garden', 'mat', 'oven', 'roof' 'thatch', 'spear', 'village'	9
	Verbs / Adjectives		
C1	Motion / Posture	'(be) in/on', 'come', 'sit', 'stand', 'crouch'	5
C2	Basic Actions / Events	'bathe', 'count', 'die', 'eat/drink', 'give', 'hear', 'hold', 'laugh', 'recline', 'pierce', 'plant', 'search', 'sing', 'throw', 'walk', 'yawn'	16
C3	Basic States	'bad', 'black', 'dry', 'far', 'good', 'itchy', 'new', 'short', 'slippery', 'thick', 'white', 'yellow', 'young'	13
C4	Time and Location	'right', 'yesterday'	2
D	**Numerals**	'one', 'two', 'five', 'six', 'tens'	5

spondence sets (distinguished below with subscript numerals). The supporting data for each of these sets is provided in the course of demonstrating the correspondences in section 4.

4. SOUND CORRESPONDENCES. In this section, we describe the 35 consonant correspondences that we have identified in our sample of AP languages. In most cases, the correspondences are conditioned by environment; we thus provide examples of the correspondences in word-initial, word-medial, and word-final position. Tables 3–8 set out the consonant correspondences identified in our sample of AP languages, as well as the reconstructed Proto–Alor-Pantar (PAP) phoneme for each correspondence set. The environment (Env) column indicates whether the correspondence applies in initial (#__), medial (V__V), or final (__#) position. A zero (Ø) in a column indicates that the PAP sound in question is lost in that language. An empty box in a column indicates that we lack sufficient data to posit a reflex for that language. A slash indicates that more than one reflex is found in that language.

Transcription follows IPA conventions.[10] Geminate consonants and long vowels are indicated with a length mark (:). Word stress is transcribed here only where relevant to the correspondence in question (for example, 'dog' in table 42). In most of the modern languages, stress is on the penultimate syllable; however, stress may also be attracted to heavy syllables, as in TWA *ji'var* 'dog'. In addition, stress may be phonemically contrastive in some languages, as in WP *ba'wa* 'conch shell' vs. *'bawa* 'drum'.

In the tables, the languages are arranged in order roughly from west to east, with the westernmost languages on the left and the easternmost languages on the right. This arrangement is maintained throughout all the tables in the paper.

In the following, we discuss the correspondences in word-initial, word-medial, and word-final position separately for each consonant. By examining the correspondences in each position separately, we are able to tease out apparent or false cognates that show the expected form in initial position but an unexpected reflex in medial or final position. Nevertheless, such irregular forms may be included in correspondence sets when they serve to demonstrate the correspondence under discussion. In these cases, the irregular forms are denoted with a preceding double dagger (‡). For some of these forms, we can identify the form as borrowed from a particular source language, but for many the reason for the irregularity has not yet been identified. Finally, in the tables below, we reconstruct PAP forms

TABLE 3. ALOR-PANTAR VOICED STOP CORRESPONDENCES

PAP	Env	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
*b	#_	b	b	b	b	b	b	b	b	f	p	p	p
*b	V_V	f/v	f/v	b	b:	b	b	b	b	b	f	p	p
*b	_#	f/v	f/v	b	p	b	b	b	b	Ø	p	p	p
*d	#_	d	d	d	d	d	d	d	d	r	t	d	d
*d	V_V	d	d	d	d:	d	d	d	d	r	t	d	d
*d	_#	r	r	d	r	d	d	d	r	r	t	d	d
*g	#_	g	g	g	g	ʔ	ʔ	g	g	h	g	g	g
*g	V_V	ɦ	x	g	g:	Ø	ʔ	g	g	h	Ø	j	l

10. The IPA transcriptions used in this paper differ from the Indonesian-based orthographies of Alor-Pantar languages we use in other publications. Important differences include IPA /j/ = orthographic *y,* /tʃ/ = *c*, /dʒ / = *j*.

only when we have broad geographic evidence. That is, reflexes must be found in minimally one language of Pantar (TWA, NDB, KER, WP), one language of West Alor and the Pantar Strait (BLG, ADN, KLN, Kui), and one language of East Alor (ABU, KMN, SWL, WRS). Where reflexes are found only in a restricted region such as Pantar or Eastern Alor, we do not reconstruct a PAP lexeme.

TABLE 4. ALOR-PANTAR VOICELESS STOP CORRESPONDENCES

PAP	Env	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
*p	#_	p	p	p	p	p	p	p	p	p	f	p	p
*p	V_V	p	p/f	p	p:	p	p		p	p	f		
*t	#_	t	t	t	t	t	t	t	t	t	t	t	t
*t	V_V	t	t	t	t:	t	t	t	t	t	t	t	t
*t	_#	t	t	t	t	t	ʔ	t	t	t	t	t	t
*k	#_	k	k	k	k	k	ʔ	k	k	k	k	k	k
*k	V_V		k	k	k:	Ø	ʔ	k	k	k	k	k	k
*k	_#	k	k	k	k	Ø	Ø	k	k	k	k		Ø
*q	#_	q	q	x	g	k/ʔ	ʔ	k	k	k	k	k	k
*q	V_V	q	q	x	k	Ø	Ø	k	k	k	k	k	k

TABLE 5. ALOR-PANTAR FRICATIVE CORRESPONDENCES

PAP	Env	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
*s	#_	s	s	s	s	h	h	h	s	t	s	t	t
*s	V_V	s	s/tʃ	s	s	s	h	h	s	t	s	t	t
*s	_#	s	s	s	s	Ø	h	h	s	t	h	t	t
*h	#_	h/ħ	Ø	Ø	h	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø

TABLE 6. ALOR-PANTAR NASAL CORRESPONDENCES

PAP	Env	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
*m	#_	m	m	m	m	m	m	m	m	m	m	m	m
*m	V_V	m	m	m	m:	m	m	m	m	m	m	m	m
*m	_#	m	m	m	Ø	ŋ	ŋ	n	n	m	m	m	m
*n	#_	n	n	n	n	n	n	n	n	n	n	n	n
*n	V_V		n	n	n:	n	n	n	n	n	n	n	n
*n	_#	n	ŋ	ŋ	ŋ	ŋ	ŋ	n	n	ŋ	ŋ	ŋ	ŋ

TABLE 7. ALOR-PANTAR LIQUID CORRESPONDENCES

PAP	Env	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
*l	#_	l	l	l	l	l	l	l	l	l	l	l	l
*l	V_V	l	l	l	l	l	l	l	l	l	l	l	l
*l	_#	i	Ø	i		l	i	l	l	Ø	i	l	l
*r	V_V	r	l	r	l	r	l	r	r	j	l	r	r
*r	_#	r	Ø	r	Ø	r	i	r	r	i	i	r	r

TABLE 8. ALOR-PANTAR GLIDE CORRESPONDENCES

PAP	Env	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
*w	#_	w	w	w	w	v	w	w	w	w	w	w	w
*j	#_	j	j	j	j	dʒ	s	Ø	j	j	j	j	j

4.1 VOICED STOPS. We reconstruct three voiced stops in labial, alveolar, and velar positions. Labial and alveolar voiced stops are well attested in all positions. However, the evidence for a voiced velar stop in initial position is based entirely on the observation of third person pronominal forms. The reflexes of initial PAP *b are shown in table 9. In the Central Alor language ABU, *b is regularly reflected as a fricative /f/, whereas in Eastern Alor languages KMN, SWL, and WRS, it is devoiced as /p/. Everywhere else *b is retained as /b/.

We do not reconstruct a word for 'maize' to PAP, but it is included here because its consonant correspondences follow the established patterns. Maize was first introduced into the region in the fifteenth to sixteenth century. AP lexemes for 'maize' represent indirect borrowings of Old Malay *batari* 'sorghum', which diffused across the languages as the crop spread. Since the historical record indicates that maize was first introduced into agriculture into western Timor, it is most likely that Austronesian languages of Timor were the source for 'maize' lexemes in AP (for example, Tetun *batar* 'maize').

Similar issues of borrowing surround the reconstruction of 'betel nut' in PAP. The betel or areca palm (*Areca catechu*) is known to have been domesticated in mainland Southeast Asia (Yen 1977). However, there is no archaeological evidence as to when the domesticated palm would have reached the Alor archipelago. There is linguistic and archaeological evidence that Proto-Austronesians in Taiwan had betel (that is, 'betel' is reconstructible to Proto-Austronesian), and that Austronesians transported betel at some points in their dispersal (Lichtenberk 1998). The similarity of the AP lexemes for 'betel' and those in surrounding Austronesian languages (for example, Tetun *bua* 'betel', Tokodede *buo* 'betel') suggests that AP 'betel' lexemes may in fact be borrowings from Austronesian. Given this lexical likeness and the uncertainity of the timing of the arrival of betel in the region, we tentatively reconstruct a PAP (loan) lexeme for 'betel nut'.

The medial reflexes of *b are shown in 0. One of the most striking features of this correspondence set is the presence of gemination in WP. The gemination process is a characteristic feature of WP; most PAP stops (including nasal stops) are geminated in medial position in WP (transcribed here as long consonants *b:*, *d:*, etc.). In particular, we infer that modern nongeminate medial stops in WP reflect either an original consonant-final form or borrowing from another AP language. In modern WP, there is a robust phonemic contrast between geminate and nongeminate consonants, as between *duba* 'slippery' and *dub:a* 'push.' Phonetic geminates do occur in some other AP languages, notably Nedebang and Sawila; however, there is little evidence that geminates have phonemic status in those languages. Furthermore, only in WP do we find geminates as a regular reflex of PAP medial stops; elsewhere they occur sporadically.

TABLE 9. INITIAL *b

Gloss	PAP *b	TWA b	NDB b	KER b	WP b	BLG b	ADN b	KLN b	Kui b	ABU f	KMN p	SWL p	WRS p
pig	*baj	bai	bei	bei	bai	be	bi	beːʔ	bei	fe	pe	pi	pei
betel nut	*bui	bui	buja	bui	bu	bu	buː	bɔiʔɪh	bui	fu		pu	pui
axe	*balin				baliŋ		baliŋ*			faliŋ	paliŋ		
maize	—	batar	baːta	batar	batːe	batar	batiʔ	bat	batar	fati	patei	patar	peter
crocodile	*bagai	baɦai		bagai	‡bagai		baʔai	bəgɑi	‡buai	fahai	piee		

* Denotes 'shovel'.

TABLE 10. MEDIAL *b

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*b	ɸ/v	f/v	b	b:	b	b
village	*haban	haɸan	‡afaŋ	aban	hab:aŋ	abaŋ	baŋ
dog[1]	—	jivar	bar	ibar	jab:e	dʒabar	‡bɛl
spear	*qaba(k)	‡qab	‡qaba	xabi	kab:i	‡ʔaba	boko
star	*jibC	jiɸ	jifa	‡ip(alaq)	hib:i		i:b
new	*siba	‡sib	sava(ʔa)	sib-	sab:a*	hiba	‡habar

Gloss	PAP	KLN	Kui	ABU	KMN	SWL	WRS
	*b	b	b	f	p	p	p
village	*haban	eben	aban	afɛŋ†			
dog[1]	—						
spear	*qaba(k)	kəbɑk	kabak	kafak	kapa		
star	*jibC	ib	ib(ra)				
new	*siba	həbɑ:	saba	tɪfɑ	supa(ka)	tipea	təpa

* Denotes ‘new sprout’.
† Denotes ‘small settlement, hamlet (n.)’, ‘dwell, live in a place, settle (v.)’.

Distinguishing medial and final correspondence sets can be difficult, since there are many instances where medial consonants in some languages correspond to initial or final consonants in other languages. For example, NDB *bar* ‘dog’ lacks the initial syllable found in cognate forms and hence retains the stop reflex rather than the fricative reflex expected in medial position.

The correspondence sets for final *b include modern reflexes that are medial, due to the presence of an epenthetic vowel. Thus, ABU *afu* ‘fish’ and *-lifi* ‘tongue’ behave as medial consonants, reflecting *b as a fricative and indicating that epenthesis must have occurred prior to the weakening of *b. On the other hand, the fact that *b in WP *-lebu* ‘tongue’ does not geminate indicates that epenthesis in this case must have occurred after the gemination process. The initial hyphen in these forms indicates that they are obligatorily possessed and must be preceded by a pronominal prefix.

The variation in TWA and NDB between voiced and voiceless reflexes of noninitial *b appears to be unconditioned. Klamer (2010:38) notes that while /ɸ/ and /v/ are distinct phonemes in TWA, the voiced variant is quite rare. The sporadic voicing seen in these correspondence sets may reflect a recent phonemicization of /v/. At this time, we are unable to determine what conditioning, if any, determines the TWA and NDB reflexes of noninitial *b.

The correspondences for initial *d are given in 2. ABU and KMN have /r/ and /t/, respectively, while the other languages retain /d/.

The correspondences supporting medial *d are given in 3. ABU and KMN have /r/ and /t/, respectively, as they do for the initials. In WP, *d geminates as expected, while the

TABLE 11. FINAL *b

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*b	ɸ/v	f/v	b	p	b	b	b	b	Ø	p	p	p
wave	*bob	boːɸ	bova	boːb			bobo	boːb		fɔ			
fish	*hab(i)	ɦaɸ	aːfi	ab	hap	aːb	aːb	əbi	eb	afu	api	api	api
tongue	*-leb(ur)	-livi	-lefu	-leːb	‡-lebu	‡-dʒebur	-lɛb	-lɛb	-liber	-lifi	‡-opei		‡-jebur

remaining languages retain *d*. The irregular correspondences for 'bat' in NDB *mara* and KER *merei* may reflect borrowing. Alternatively, these forms may have a more complex history in which final syllables were originally lost, leading these forms to be treated as final (see 4). In ADN, *d was palatalized following a vowel sequence ending in a high front vowel (for example, *mudi > muid > mudʒ 'body hair'). This has led to a phonemic distinction between alveolars and palato-alveolars (/d/ and /dʒ/, /n/ and /ɲ/) in some dialects of modern Adang (Haan 2001). KLN has nonphonemic palatalization in the same environment, while the closely related language Kabola does not undergo palatalization (cf. Kabola *muir* 'body hair').

As shown in 4, the correspondences for final *d are almost identical to those for medial *d, except that TWA, NDB, WP, and Kui reflect *d as *r* in this environment. The KER form *wer* 'sun' is likely a borrowing from neighboring TWA or NDB.

While PAP *g is robustly attested in noninitial environments, the reconstruction of initial *g hinges entirely on the correspondence of third person prefixal forms in PAP, as shown in 5. That is, all instances of initial /g/ in modern AP languages can be traced to third person pronouns. A correspondence set reflecting an original bound third person singular prefix *ga- has reflexes in all languages and can be reconstructed for the protolanguage. This prefix has reflexes with initial *g* in all languages except Blagar and Adang, which have glottal stop, and ABU, which has a glottal fricative. A third person plural counterpart to this prefix is attested in a few languages. This form can be tentatively reconstructed as *gi-. It is absent in modern AP languages such as ADN, KLN, KMN, and ABU, which have generalized their reflexes of the PAP third person singular prefix to both singular and plural contexts. A third reflex of initial *g is found in the third person

TABLE 12. INITIAL *d

Gloss	PAP *d	TWA d	NDB d	KER d	WP d	BLG d	ADN d	KLN d	Kui d	ABU r	KMN t	SWL d	WRS d
rat	*dur	dur	dur	dur	di	duru	du:l	dɔr	dur	rui	tui	daru	dur(ki)
sing	*dar(a)	da:r			dali	dar	da:la		dar			dara	dra
bird	*dVl	dei	daja			‡ul			adol	ruo	atoi	adala	adol
slippery	*dul(a)			duj-	‡duba	dula	dulu?	du:l	dula	rula	tula(ka)		dol(ok)

TABLE 13. MEDIAL *d

Gloss	PAP *d	TWA d	NDB d	KER d	WP d:	BLG d	ADN d
plant (v.)	*mudin	midan		muduŋ	mid:iŋ	mudiŋ	mudiŋ
bat	*madel	mədi	‡mara	‡merei	mad:e	demel*	
right	—	jidan	jediŋ		jad:iŋ		
body hair	*mudi	-mud	mudi	-mudu?		-mudi	mudʒ

Gloss	PAP *d	KLN d	Kui d	ABU r	KMN t	SWL d	WRS d
plant (v.)	*mudin	mdin	medi	murow	mit	madiŋ	mdi
bat	*madel		madel	mare	matei	madi:(ku)	
right	—						
body hair	*mudi	amudi		amuri		madi	-mudi

* This form is metathesized.

TABLE 14. FINAL *d

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*d	r	r	d	r	d	d
sun	*wad(i)	war (get)	weri	‡wer	war*	ved	fɛd
throw	*-od	‡os		-oːd		oda	ɔd
fire	*had(a)	ħar	ar	aːd		aːd	
garden	—	maħar	maxara		mag:ar		maʔad
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	*d	d	r	r	t	d	d
sun	*wad(i)			wari	wati	wadi	widi
throw	*-od	-ɔːd	-or	-uor	wota†		
fire	*had(a)	ədɑ	ar	ara	ati	ada	ada
garden	—						

* Denotes 'shine, burn' (cf. *was* 'sun').
† Denotes 'beat, strike (drum)'.

genitive marker *ge(-), which indexes alienable possessors (in contrast to *ga-, which indexes inalienable possessors). The reconstruction of genitive *ge(-) is supported by the presence of reflexes in a robust geographical spread of AP languages. A final correspondence set supporting *g is found in the third person locative prefix in several languages of Alor. There is no evidence for this prefix in the languages of Pantar (TWA, NDB, KER, WP, BLG), and we do not reconstruct it to PAP. Note that KMN has a regular change of initial *g to /w/ before back vowels, hence the form *wo-*.

With some possible exceptions, the forms cited in 5 are bound, occurring as prefixes with either nominal or verbal roots. Exceptions include Adang *ʔe* and Klon *ge* 3GEN.[11] At this stage, we remain agnostic as to whether the PAP genitive was a free or bound form. Other free pronouns vary in their form across the modern AP languages and cannot be reconstructed to PAP (Kratochvíl et. al. 2011).

In medial position, a correspondence set reflecting *g is evidenced in numerous lexical forms, as shown in 6. Only in KER, KLN, and Kui is *g retained unchanged in medial position. In WP we find the expected geminate in all forms except *bagai* 'crocodile', which may be a borrowing from KER. In TWA, medial *g is reflected as a pharyngeal fricative, while in Abui it is reflected as a glottal fricative. Other languages reflect either a glottal stop, a liquid, a fricative, or zero. However, medial reflexes in SWL and WRS are supported by

TABLE 15. 3RD PERSON PRONOMINAL FORMS REFLECTING INITIAL *g

Type	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*g	g	g	g	g	ʔ	ʔ	g	g	h	g	g	g
3SG	*ga-	ga-	ga-	gV-*	ga-	ʔa-	ʔa-	g	ga-	ha-	ga-	ga-	gV-
3PL	*gi-	gi-			gi-	ʔi-						gi-	gi-
3GEN	*ge(-)			ga-	gai-	ʔe-	ʔe†	ge		he-	ge-	ge-	
3LOC	—						ʔo-	go-		ho-	wo-		

* Prefix vowel harmonizes with stem vowel.
† In Adang, *ʔe* has been restricted to marking possessors in contrastive focus.

11. The Klon form is analyzed by Baird (2008) as a free form based on its ability to occur following an NP. Yet it is equally possible that Klon has homophonous bound and free genitive forms differing in distributional restrictions, analogous to WP *gai-* (bound) and *ga'ai* (free).

TABLE 16. MEDIAL *g

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*g	h	x	g	g:	Ø	ʔ
yellow	—	baħari	baxori	bagari	bug:a	bori	baʔoi
yawn	—	ħaħar		agur			
banana	*mogol	moħoi		mogoi	mag:i	mo:l	moʔoi
garden	—	maħar	maxara		mag:ar		maʔad
crocodile	*bagai	baħai		bagai	‡bagai		baʔai
hear	—					me	

Gloss	PAP	KLN	Kui	ABU	KMN	SWL	WRS
	*g	g	g	h	Ø	j	l
yellow	—	(bʊ)bʊgɔr*	bagura				
yawn	—		agu	ahau			
banana	*mogol	mɔgɔl			mo:i		mulul
garden	—						
crocodile	*bagai	bəgɑi	‡buai	fahai	piee		
hear	—	məgɪh	magi	mahia	-mai	maji:ŋ	

* The Klon form appears to have fossilized with a reduplicated CV.

only one lexical item. It is striking how few of the medial *g lexical items can be reconstructed to PAP because they do not occur across a range of languages.

The evidence for *g in final position is extremely weak. In the modern languages, final *g* occurs only in the Pantar languages TWA and KER (as well as Sar, not in our sample). In our 400-item wordlist, final *g* is found in only eleven distinct TWA word forms. None of these has cognates in a central or eastern Alor language. Cognates with Pantar and western Alor languages do exist; however, in many cases the correspondence is between medial *g* and final *g*. For example, TWA *miaag*, NDB *miaagi*, KER *miag* 'yesterday'; and TWA *bog*, NDB *boga*, WP *bog:a* 'young'. Hence, it seems plausible to conclude that TWA and KER final *g* actually derive from medial *g, and that PAP *g was not found in final position.

4.2 VOICELESS STOPS. As shown in 7, PAP *p remains unchanged in all the languages except KMN, where it weakens to *f*, and WP, where it is geminate in medial position. SWL and WRS merge *b and *p as *p*. Note that WP *par* 'scorpion' must be a loan from a language that preserves final *r (cf. table 40 below).

While *p* is found in final position in some modern languages, we find no evidence to support reconstruction of *p in final position. Rather, final *p* results from loss of final vowels (for example, *tapai > TWA *tap* 'pierce') or as a reflex of *b (for example, *hab(i) > WP *hap* 'fish').

There is a regular and unchanging correspondence of initial *t* across all the AP languages (as shown in 8), thus securing the reconstruction of PAP initial *t.

In medial position, *t is also retained in all languages except WP, which has a geminate medially; see 9. Reflexes of the forms meaning 'dry' and 'oven' are not sufficiently widely distributed to justify reconstruction at the level of PAP. As stated earlier, we don't reconstruct PAP 'maize' since it is known to be a late borrowing from Austronesian. The resemblance between PAP *-tan 'hand/arm' and Malay *taŋan* 'hand' is superficial only and cannot be taken to indicate that the AP lexemes are Austronesian borrowings. The

form *taŋan* for 'hand' is a lexical innovation of Malayic and cannot be reconstructed to higher levels of the Austronesian family: Proto-Austronesian and Proto–Malayo-Polynesian reconstructions are *lima for 'hand' and *baRa for 'arm', and it is reflexes of these

TABLE 17. INITIAL AND MEDIAL *p

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*p	p	p	p	p/p:	p	p
hold*	*p(i,u)nV	pin	pini	pin-	pin:i	pina	pun
1PL.INCL	*pi-	pi-	pi-	pi-	pi-	pi-	pi-
scorpion	*pVr	par		par	‡par	per	pai
pierce†	*tapai	tap	tapa	tap-	tap:a(ŋ)	tapa	ta:pa(ŋ)
search	—			rap-		rapiŋ	la:p
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	*p	p	p	p	f	p	p
hold†	*p(i,u)nV	pʊɪn	puna	puna	fun	puni‡	poiŋ**
1PL.INCL	*pi-	pi-	pi-	pi-		pi-	
scorpion	*pVr	par	per	pei	‡fal		per(buk)
pierce‡	*tapai	tap	təpa(n)	tapai	tapei	tafe	
search	—			-rap			

* Reflexes of *p(i,u)nV typically encompass the meanings 'hold' and 'grab', with the difference depending on the prefixation of the verb.
† Reflexes of *tapai encompass the meanings 'pierce', 'stab', 'plant in the ground', and 'pound rice'.
‡ SWL has *wuni* 'hold' and *puni* 'hit'.
** WRS has *woiŋ* 'hold' and *poiŋ* 'hit'.

TABLE 18. INITIAL *t

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*t	t	t	t	t	t	t
recline	*tia	ti:ʔ	taʔa	te:	tiʔaŋ	tiaʔ	
saltwater	*tam	‡taʔ		tam	tawa	taŋ	taŋ
short	*tukV	tuk	tuku	tuk-	tuk:a	‡tuka(ŋ)	toʔo(ŋ)
stand	*tas	tas	tasi	tas-		tai	tohɔ
tree	*tei	tei	tei	tei		te	ti
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	*t	t	t	t	t	t	t
recline	*tia	tɑ:	ta	taha	ta:*	(ga)te	taj
saltwater	*tam	tan	tan	tama	tama	tama	tamaʔ
short	*tukV	tuk	tuk	tuku	tuk†	tuku(da)	tuk
stand	*tas	(mə)tɛh		(na)tet			-tati
tree	*tei	(ə)təʔ	(a)tei	(ba)taa			

* Denotes 'on top'.
† Denotes 'short piece, cutting'.

TABLE 19. MEDIAL *t

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*t	t	t	t	t:	t	t	t	t	t	t	t	t
dry	—					tata	taʔata	təkɑt	takata	takata			
maize	—	batar	ba:ta	batar	bat:e	batar	batiʔ	bat	batar	fati	patei	patar	peter
oven	—	‡tutah	tutu-	tutuk	tut:u	tutu							
hand/arm	*-tan	-tan	-taŋ	-taŋ	-t:aŋ	-taŋ	-taŋ	-tɑn	-tan	-taŋ	-taŋ	-taŋ	-teŋ

protolexemes for 'hand' and 'arm' that are found in the Austronesian languages surrounding the AP languages. Malay has only been present in the region since the historical period, and Malay influence on the AP languages might have started as late as the beginning of the twentieth century.[12] As such, we unproblematically reconstruct PAP *-tan 'hand/arm'.

Final *t is preserved in only a few of the modern languages. Forms for 'betel vine' are problematic, as it has a medial *t* in more than half of the modern languages. We tentatively reconstruct this form with final *t based on two pieces of evidence. First, the WP form *meta* does not reflect gemination, which would be expected as a reflex of medial *t. Second, several of the languages have a long vowel or diphthong. We thus reconstruct *mait and presume a process of palatalization following a high front vowel. Thus, *t > *tʃ* / Vi_ in ADN (see 4.1), *t > *h* /Vi_ in KLN (presumably via [s]), and *t > *s*/ Vi_ in Kui, KMN, SWL, and WRS. In table 20 we list only the original reflex, not the secondary development reflected in 'betel vine'. However, we note that the irregularity of the modern forms points to it being a borrowing; see also the case of 'betel nut' in 4.1 above.

There is a robust correspondence of initial *k* across the AP languages, which secures the reconstruction of PAP *k. Multiple cognate sets consistently reflect *k as /k/ in all modern languages except for ADN and BLG, as illustrated in table 21. In ADN, *k is reflected as glottal stop, and in BLG it is lost. While BLG *ira* 'comb' lacks initial *k*, evidence from other Blagar dialects shows that Proto-Blagar did indeed have initial *k* in this form, supporting PAP *k > BLG *k* (Robinson 2011). Note that the medial correspondences for NDB 'comb' and ABU 'fingernail' are irregular, and Kui 'fingernail' has an unexpected final velar nasal.

The noninitial correspondences of *k in table 22 are very similar to the initial correspondences, except that WP shows gemination. The lexeme 'lizard' is likely an Austronesian

TABLE 20. FINAL *t

Gloss	**PAP**	**TWA**	**NDB**	**KER**	**WP**	**BLG**	**ADN**
	***t**	**t**	**t**	**t**	**t**	**t**	**ʔ**
coconut	*wat(a)	wat	wata	wa:t	wata	vet	faʔ
betel vine	*mait	met	mata	ma:t	meta	ma:t	metʃe
leg	—	-fat		-bat	-uta		
flea	*kVt	‡hat			kati	atuk	
wound	—	bat	bata	bat		bata	
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	***t**	**t**	**t**	**t**	**t**	**t**	**t**
coconut	*wat(a)		‡bat	wata	wate	wata	wata
betel vine	*mait	mɛh	mesin	metiŋ	maisi	ma:si	mas
leg	—						
flea	*kVt		kot	(tu)kota			tokoʔ
wound	—		bata				

12. It is likely that Malay was only introduced to the Papuan speakers on Alor and Pantar through the Dutch schools that were opened in the early twentieth century. For example, Du Bois (1944:223) notes that among the Abui people with whom she lived in the 1930s, Malay was only known by school children. The first Dutch schools were opened on Alor in 1906, and on Pantar in the 1920s (Klamer 2010:14). In the Dutch schools, the language of education was Malay, as elsewhere in the Dutch East Indies.

borrowing (cf. Alorese *take*), perhaps explaining the anomalous reflexes ADN *teko* and KMN *tak:e:*. However, this form is geminated as expected in WP *tak:e*.

Final *k is retained in all languages except BLG, ADN, SWL, and WRS, where it is lost, as shown in table 23. None of our correspondence sets have cognates in Sawila; however, final *k* is rare in our Sawila data set, occurring in only two forms: *werpa:k* 'frog' and *kispa:k* 'earthworm' (both lexical innovations shared with WRS). The correspondences for final *k can be difficult to tease apart from those for medial *k, since many languages reflect later vowel epenthesis or apocope. We take the presence of a geminate in WP to be diagnostic in this regard, since WP geminates do not occur word-finally. This criterion is admittedly problematic, since it is entirely possible that vowel epenthesis preceded gemination in WP. Furthermore, WP sometimes lacks cognates for relevant lexical items, as with 'horn' in table 23. The BLG form *-muʔu* ' horn' may indicate that this form also reflects an original medial rather than final stop, further weakening the case for final *k.

The reconstruction of *q is supported by the presence of a postcoronal voiceless obstruent phoneme distinct from the velar stop in several Pantar languages. In TWA and NDB, this is a uvular stop *q*; in KER, it is a velar fricative *x*. KER *kiŋ* 'mosquito' is an

TABLE 21. INITIAL *k

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*k	k	k	k	k	k	ʔ	k	k	k	k	k	k
comb	*kir	kir	‡kiri	kir	kiri*	kiri	‡ʔil	kir	kir(mei)	‡kir		(ke:)kuri	kuri
bone	—	kir	kili	kiri		‡ira							
dog	—							kur	kur	ka:i	kui		
finger nail	*kusin		kutʃiŋ	kusiŋ	kusi			kuh	‡kusɪŋ	‡kusɪŋ	kuisiŋ		

* Denotes 'fork.'

TABLE 22. MEDIAL *k

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*k	k	k	k	k:	k	ʔ
crouch	*luk(V)				luk:iŋ		
short	*tukV	tuk	tuku	tuk-	tuk:a	‡tukaŋ	toʔoŋ
good	—						noʔo
lizard	—	takok	takara:b	tek	tak:e	teˈke	‡teko

Gloss	PAP	KLN	Kui	ABU	KMN	SWL	WRS
	*k	k	k	k	k	k	k
crouch	*luk(V)		luk	luk	luk*		lukuk
short	*tukV	tuk	tuk			tuku(da)	tuk
good	—	nok	noka				
lizard	—	takek	taˈkok	tekok	‡tak:e:		

* Denotes 'bow, bend'.

TABLE 23. FINAL *k

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*k	k	k	k	k	Ø	Ø	k	k	k	k		Ø
one	*nuk	nuk	nuku	nuk-	anuku	nu	nu	nɔk	nuku	nuku	nok		‡no
horn	*-muk			-muk		-mu	‡-muʔu	-mɔk	-muk	-muk	‡-mu		

exception that may reflect influence from BLG. BLG itself shows alternation between a velar and glottal reflex of *q. Note that the *r* in 'tens' behaves as a medial consonant, since this numeral formative only occurs in compounds with following numeral, as with TWA *qar nuk* 'ten' (see table 39). Initial reflexes of *q are given in table 24.

The medial reflexes of *q, shown in table 25, are similar to those in initial position, except that both BLG and ADN show loss of medial *q. BLG *kaka* and ADN *kakaʔ* 'itchy' are anomalous, as they retain the medial consonant. BLG *mədʒa* 'white' is, in fact, cognate due to a regular process of glide insertion between the vowels /i/ and /a/, followed by glide fortition: *miaqa > *mia* > *mija* > *mədʒa*. The most interesting reflex of medial *q is found in WP. Unlike the other voiceless stops, the uvular stop is not reflected as a geminate in WP, but instead as a nongeminate velar stop.

Hence, *q provides an additional source for nongeminate intervocalic voiceless velar stops in WP. This, in turn, may inform reconstruction of final vowels in PAP. Since *q does not geminate in WP, WP *alaku* 'two' can readily be derived from *raqu, supporting reconstruction of the final vowel. On the other hand, WP *anuku* 'one' corresponds to TWA and NDB forms with velar stops, hence the reconstruction of PAP 'one' must contain a velar, not a uvular. The fact that WP *anuku* does not contain a geminate means that either it has been borrowed or that the vowel has been added following the gemination

TABLE 24. INITIAL *q

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*q	q	q	x	k	k/ʔ	ʔ
spear	*qaba(k)	‡qab	‡qaba	xabi	kab:i	‡ʔaba	boko
itchy	—	qa:q	qaqa	xax-	kaka	‡kaka	‡kakaʔ
mosquito	*qin	qiʔin	‡kim(balu)	‡kiŋ	ki*	kinit	ʔin
tens	*qar-	qar-	qa-	xar-	ke-	ʔar-	air-
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	***q**	**k**	**k**	**k**	**k**	**k**	**k**
spear	*qaba(k)	kəbɑk	kabak	kafak	kapa		
itchy	—	ka:k					
mosquito	*qin	ikin	kin	kuim(a)	kiŋ(ba)	ka(we:ŋ)	ku(buŋ)
tens	*qar-	kar-	kar-	kar-	kar-		

* Denotes 'maggot'.

TABLE 25. MEDIAL *q

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*q	q	q	x	k	Ø	Ø
two	*araqu	raq	raqu	rax-	alaku	aru	a:lu
itchy	—	qa:q	qaqa	xax-	kaka	‡kaka	‡kakaʔ
white	—	miaq	miaqa	miex-	miaka	mədʒa	
black	*aqana	qaʔan	qana	xan-	‡ana	kana	(la)ʔanaʔ
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	***q**	**k**	**k**	**k**	**k**	**k**	**k**
two	*araqu	ərʊk	oruku	‡ajoku	‡ok	‡jaku	‡joku
itchy	—	ka:k					
white	—						
black	*aqana	(ɑʔ)ɑkɑn	akana	akani		akana	akeŋ

process. In the absence of any evidence for borrowing, we reconstruct *nuk 'one' without a final vowel (see table 23).

The evidence for *q in final position is extremely limited. One possible example is KER *banax* 'smoke', which corresponds to WRS *punak*, with final velar stop. However, the TWA and NDB forms *bu:n* and *bun*, respectively, lack the expected final uvular. Moreover, this form could well be an Austronesian loan (cf. Proto-Austronesian *CebuN). Another candidate correspondence is 'rice': WP *ala* and KLN, Kui, WRS *arak*, which compare to TWA *qar*, NDB *qara*, and KER *(na)xar*. If the TWA, NDB, and KER forms are interpreted as a result of metathesis of *r and *q, then this correspondence could also support *q in final position, namely, *araq.

4.3 FRICATIVES. We reconstruct two fricatives in PAP, *s and *h. While *s occurs freely in all positions, the glottal fricative *h is restricted to initial position. Correspondence sets for *s are relatively straightforward. In initial position *s weakens to *h* in BLG, ADN, and KLN, and strengthens to *t* in ABU, SWL, and WRS, as shown in table 26.

The medial correspondences for *s are almost the same as the initial, except that BLG retains *s* and NDB sometimes has an affricate *tʃ*. The evidence for the medial reflex of *s in Blagar is weak, since the two Blagar forms in table 27 exhibit irregular correspondences in other positions, and hence may be loans. The correspondences for final *s in table 28 differ only in that final *s is lost in BLG and retained as a fricative rather than an affricate in NDB. NDB *bi:* 'mat' is an exception in that it lacks final *s*.

The reflexes of *h are given in table 29. WP retains *h while all the remaining languages, except TWA, lose *h. Note that TWA has two reflexes of *h, the glottal fricative *h* and the pharyngeal fricative *ħ*. This is due to a phonemic split in TWA, where *ħ* originally occurred only before back vowels, and *h* occurred elsewhere. Modern TWA still tends this way, with pharyngeal *ħ* generally occurring before back vowels and glottal *h* preceding

TABLE 26. INITIAL *s

Gloss	PAP *s	TWA s	NDB s	KER s	WP s	BLG h	ADN h	KLN h	Kui s	ABU t	KMN s	SWL t	WRS t
new	siba	‡sib	savaʔa	sib-	sabia	hiba	habar	həbɑ:	saba	tɪfɑ	supa(ka)	tipea	təpa
wind	—						hamɔi			timoi	sumui	tamuro	
shark	—	sifar	sifi	sibar	sib:u	hibir			sobor				

TABLE 27. MEDIAL *s

Gloss	PAP *s	TWA s	NDB s/tʃ	KER s	WP s	BLG s	ADN h
bad	*jasi	jas	jetʃi	jas-	jasa	‡dʒasi	saha*
fingernail	*kusin		kutʃiŋ	kusiŋ	kusi		
five	*jiwesin	jusan	jisin	isin	jasiŋ	‡jisiŋ	iwɛhɛŋ

Gloss	PAP *s	KLN h	Kui s	ABU t	KMN s	SWL t	WRS t
bad	*jasi					jaati	
fingernail	*kusin	kuh	kusɪŋ	‡kusɪŋ	kuisiŋ		
five	*jiwesin	əwəh	jesan	jetiŋ	iwesiŋ	jo:tiŋ	‡wetiŋ

* Denotes 'broken'.

TABLE 28. FINAL *s

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*s	s	s	s	s	Ø	h
sit	*mis	mis	misi	mis-	mis(iŋ)	miŋ	mihiʔ
stand	*tas	tas	tasi	tas-		tai	tohɔ
teeth	*-uas	-usan	-usiŋ	-uasiŋ	-wasiŋ	-veŋ	-wɛhɛŋ
mat	*bis	bis	‡bi:	bis	bis	bi	bahu
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	*s	h	s	t	h	t	t
sit	*mis	mɪh	misa	miti	‡nih	miti	amit
stand	*tas	(mə)tɛh					-tati
teeth	*-uas	-wɛh	-wes	-weiti	-weh	‡-wa	‡-wesi
mat	*bis		bus	‡bu:si			bitiʔ

front vowels. Klamer (2010) lists only one example of a pharyngeal fricative preceding a front vowel, namely *ħer* 'yell, shout, chant, cry aloud'. This form is cognate with WP *horaŋ*, suggesting that the original form may have contained a back vowel, thus conditioning the TWA pharyngeal. This distinction breaks down, however, before low vowels, where a clear synchronic phonemic distinction has developed in TWA.

As noted above, PAP *h did not occur in noninitial position. In the modern languages, noninitial *h* derives from *s. In TWA, noninitial *ħ* derives from *g (see 6 above).

4.4 NASALS. There is a regular and unchanging correspondence of initial /m/ across all the AP languages. The reconstruction of PAP initial *m is thus secure and supported by multiple cognate sets, four of which are given in table 30.

In medial position, *m is retained as such except in WP, where it is geminated as part of a regular process of germination; see table 31. Subsequent changes may result in nasal-final forms that obey language-specific constraints. WP does not admit final nasals other than velars, hence WP *-haŋ* 'breast' results from later apocope, namely, PAP **hami* > *ham:* > *ham* > *haŋ*.

Final *m is retained as *m* in seven of the twelve languages, but only in TWA and KER does it occur in final position. BLG, ADN, KLN, and Kui also have final nasal reflexes, though not labial. (Kui *talama* 'six' is likely a borrowing from ABU.) The remaining languages except WP, which drops final *m, retain *m, but not in final position, due to ensuing epenthesis. Note that the TWA, KER, NDB, and KMN forms for 'six' also reflect irregular loss of *l. See table 32.

There is a regular and unchanging correspondence of initial *n in all the languages, as shown in table 33.

TABLE 29. INITIAL *h

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*h	h/h	Ø	Ø	h	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
village	*haban	hafan	afaŋ	aban	hab:aŋ	abaŋ	baŋ	eben	aban	afɛŋ			
2SG	*ha-	ha-	a-	a-	ha-	a-	a-	a-	a-	a-	a-	a-	e-
fish	*hab(i)	ħaf	a:fi	ab	hap	a:b	a:b	əbi	eb	afu	api	api	api
fire	*had(a)	ħar	ara	a:d	had:i*	a:d		ədɑ	ar	ara	ati	ada	ada

* Denotes 'to burn (of land)'.

The medial correspondences for the alveolar nasal in table 34 mirror those for the labial nasal. Medial *n is geminated in WP, but is retained as *n* in the other languages.[13]

TABLE 30. INITIAL *m

Gloss	PAP *m	TWA m	NDB m	KER m	WP m	BLG m	ADN m	KLN m	Kui m	ABU m	KMN m	SWL m	WRS m
come	*mai*	ma	ma	ma	ma	ma	ma	ma		mɛ:	me:	me	mai
betel vine	*mait	met	mata	ma:t	meta	ma:t	metʃe	mɛh	mesin	metiŋ	maisi	ma:si	mas
sit	*mis	mis	misi	mis-	misiŋ	miŋ	mihiʔ	mɪh	misa	miti		miti	amit
(be) in/on	*mi	meʔ		mi	me	mi	mi	mi	mi	mia	mi	ma	

* This reconstruction is strikingly similar to the Austronesian (Proto–Malayo-Polynesian) form **maRi* 'come', which is irregularly reflected as *ma* or *mai* in many Austronesian languages in the region; cf. Mambai (Timor) *ma*, Kambera (Sumba) *mai*. However, similar reflexes are not found in Lamaholot or Alorese, the immediate Austronesian neighbor of the Alor-Pantar languages.

TABLE 31. MEDIAL *m

Gloss	PAP *m	TWA m	NDB m	KER m	WP m:	BLG m	ADN m
fat	*tama	tamaʔ		tama		təma	tamara
bedbug	*temVk			temek		teme	
horn	*-muk			-muk		-mu	‡-muʔu
thatch	*aman	man	maŋ	maŋ		meniŋ	men
grandchild	—				‡-am:is		
thick	*dumV	‡tuʔum			dum:a		
breast	*hami	-ɦam	-ami		-haŋ	‡a:m	

Gloss	PAP *m	KLN m	Kui m	ABU m	KMN m	SWL m	WRS m
fat	*tama	təmɑd	tama	tamada			
bedbug	*temVk			tameak			mekit*
horn	*-muk	-mʊk	-muk	-muk	‡-mu		
thatch	*aman	ene:m†	amen			amaŋ	ameŋ
grandchild	—				tam	ta:mu	
thick	*dumV					dumu	dum
breast	*hami				ami	-ami	ami

* This form has metathesized.
† This form has metathesized.

TABLE 32. FINAL *m

Gloss	PAP *m	TWA m	NDB m	KER m	WP Ø	BLG ŋ	ADN ŋ
six	*talam	‡tia:m	‡tiama	‡tiam		taliŋ	ta:laŋ
saltwater	*tam	‡taʔ		tam	tawa	taŋ	taŋ

Gloss	PAP *m	KLN n	Kui n	ABU m	KMN m	SWL m	WRS m
six	*talam	təlɑn	‡talama	tala:ma	‡ta:ma		
saltwater	*tam	tan	tan	tama	tama	tama	tamaʔ

13. In certain morphological environments, gemination does not occur, for example, *gania* 'give him' vs. *gin:ia* 'giving him' (cf. Holton 2010).

Geminates are also found sporadically in NDB, reflecting either incomplete realization of the gemination development or diffusion from WP.

Final *n is reflected as a velar nasal in all languages except TWA, KLN, and Kui, where it is retained as *n*, as shown in table 35. The forms below show evidence of bor-

TABLE 33. INITIAL *n

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN	KLN	Kui	ABU	KMN	SWL	WRS
	*n	n	n	n	n	n	n	n	n	n	n	n	n
one	*nuk	nuk	nuku	nuk-	anuku	nu	nu	nɔk	nuku	nuku	nok		-no
1SG	*na-	na-	na-	na-	na-	na-	na-	na-	na-	na-	na-	na-	ne-
eat/drink*	*nai	na	na	na	na	na	naʔ	nɑːʔ	nai	ne	na	neː	nai

* Denotes 'eat' in TWA, NDB, WP, ABU, KMN; 'eat/drink' in KER, BLG, SWL; and 'drink' in ADN, KLN, Kui, WRS.

TABLE 34. MEDIAL *n

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*n	n	n	n	n:	n	n
smoke	*bunaq	buːn	bun	bənaːx	bunːa	benaŋka	banoʔo
black	*aqana	qaʔan	qana	xan-	‡ana	kana	(la)ʔana ʔ
hold	*p(i,u)nV	pin	‡pini	pin-	pinːi	pina	pun
give (to s.o.)*	*-enV	-an	-ena	‡-eŋ	-nːia	-enaŋ	-ɛn
die	*minV	min	mina	nimin-	‡hinːa	mina	miniʔ
name	*-ain(i,u)		-einu	-en	-inːu	-ene	-aniŋ
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	***n**	**n**	**n**	**n**	**n**	**n**	**n**
smoke	*bunaq	bon	bonok		puna	punaka	punak
black	*aqana	(ɑʔ)akɑn	akana	akani		akana	‡akeŋ
hold	*p(i,u)nV	pɔɪn	puna	puna	fun	puni	‡poiŋ
give (to s.o.)*	*-enV	-en	-ana		-n	-ani	-eni(r)
die	*minV		min	‡moŋ			
name	*-ain(i,u)	-ənɛʔ	-enei	-ane	-nei	-ani	

* The Timor-Alor-Pantar languages have the recipient encoded as object of the verb 'give' (Klamer and Schapper forthcoming).

TABLE 35. FINAL *n

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*n	n	ŋ	ŋ	ŋ	ŋ	ŋ
five	*jiwesin	jusan	jisin	isin	jasiŋ	‡jisiŋ	iwɛhɛŋ
hand/arm	*-tan	-tan	-taŋ	-taŋ	-tːaŋ	-taŋ	-taŋ
thatch	*aman	man	maŋ	maŋ		meniŋ	-men
fingernail	*kusin		kuciŋ	kusiŋ	kusi		
Gloss	**PAP**	**KLN**	**Kui**	**ABU**	**KMN**	**SWL**	**WRS**
	***n**	**n**	**ŋ**	**ŋ**	**ŋ**	**ŋ**	**ŋ**
five	*jiwesin	əwəh	jesan	jetiŋ	iwesiŋ	joːtiŋ	‡wetiŋ
hand/arm	*-tan	-tɑn	-tan	-taŋ	-taŋ	-taŋ	-teŋ
thatch	*aman	eneːm*	amen			amaŋ	ameŋ
fingernail	*kusin	kuh	‡kusɪŋ	‡kusɪŋ	kuisiŋ		

* This is presumably a metathesized form.

rowing among the languages, as well as possible borrowing from Austronesian. NDB and KER 'give' have final alveolar rather than the expected velar, and are likely borrowings from neighboring TWA. The correspondence set for 'fingernail' is more problematic: Kui reflects final velar instead of the expected alveolar; ABU has the velar nasal as expected but shows an irregular reflex of medial *s; and WP and KLN show irregular loss of the final nasal. It may be that this form is also a borrowing from Austronesian; however, similar forms are *not* found in neighboring Austronesian languages (cf. Alorese *tanunggul*).

4.5 LIQUIDS. We reconstruct two liquids, *l and *r, in PAP, though only *l occurs in initial position, and *r was likely not phonemically distinct in PAP (see the discussion in section 5). For expository purposes, we treat *r as if it were a phoneme in the present section. The reconstruction of initial *l is based on the cognate sets in table 36. From the present data, there appears to be a relatively regular and unchanging correspondence of initial *l* in the modern languages from which the existence of PAP *l can be posited. However, few forms are distributed widely across the languages, making it difficult to reconstruct words with initial *l.

Similarly, a regular and unchanging correspondence of *l* in medial position permits reconstruction of *l in this position as well, retained as such in all languages. A few languages show evidence of sporadic *l > *i* in medial position, for example, WRS *-iebur* 'tongue' in table 37. Also, TWA *wei* and KER *wei* 'bathe' are likely the result of loss of final vowel before the change of final *l > *i* in these laguages.

In final position, however, TWA, KER, ADN, and KMN reflect *l as *i*, as can be seen in table 38. This final vowel may be realized phonetically as a glide in the modern languages; however, we analyze these phonemically as vowels and assume the same analysis for PAP. Further, NDB, WP, and ABU lose final *l altogether.

We find insufficient evidence to reconstruct *r in initial position. In noninitial position, PAP clearly distinguished two liquids, and this distinction is preserved in most of the languages. As shown in table 39, in medial position NDB, WP, and ADN collapse *l and *r as *l* (the reflexes in KMN are less consistent). The other languages preserve *r as such. This leaves no direct historical source for *r* in NDB, WP, and ADN, and we assume that *r* in these languages has been innovated or diffused from neighboring languages. In modern WP, forms with *r* are infrequent and do not correspond regularly to other languages. In most cases they reflect lexical innovation, as in WP *re* 'bird' (compare PAP *dVl).

TABLE 36. INITIAL *l

Gloss	PAP *l	TWA l	NDB l	KER l	WP l	BLG l	ADN l	KLN l	Kui l	ABU l	KMN l	SWL l	WRS l
walk$_1$	—	lam-an*		‡amar		lamar	lamɛ	lam					
walk$_2$	—									lol	lo:	lo:la	lailol
rattan	—	liaag		le:g		lija	lɛʔ		le				
crouch	*luk(V)				luk:iŋ				luk	luk	luk		lukuk
far	—						lɛt	lɛt			letei		
monitor lizard	*lVsi		lisi				(ʔa)lɛhɛʔ			liti			

* TWA *laman* 'follow, walk along (e.g., a path)'. WP *lama* shares this sense and is likely a borrowing from TWA, which explains the lack of gemination in the WP form.

The correspondences for final *r in table 40 are similar to those in medial position, except *r > *j* in ADN and KMN, and *r > Ø in NDB and WP.[14] Irregularity in some of the correspondences may indicate that some of these sets actually reflect medial rather than final *r, for example, KMN *fal* 'scorpion' rather than the expected **fai*, and NDB *wala* 'stone' and *-ola* 'tail' rather than the expected **wa:* and **-owa.*

TABLE 37. MEDIAL *l

Gloss	PAP *l	TWA l	NDB l	KER l	WP l	BLG l	ADN l
axe	*balin				baliŋ		baliŋ*
bathe	*weli	‡wei		‡-wei		vela	wili
tongue	*-leb(ur)	-livi	-lefu	-leːb	‡-lebu	‡-dʒebur	-lɛb
sky	—	bulan		buluŋ		bulaŋ	

Gloss	PAP *l	KLN l	Kui l	ABU l	KMN l	SWL l	WRS l
axe	*balin			faliŋ	paliŋ		
bathe	*weli	-wɛl	weli	-wel	-wei	-wile	-weli
tongue	*-leb(ur)	-lɛb	-liber	-lifi	‡-opei		‡-jebur
sky	—						

* Denotes 'shovel'.

TABLE 38. FINAL *l

Gloss	PAP *l	TWA i	NDB Ø	KER i	WP Ø	BLG l	ADN i
banana	*mogol	moɦoi		mogoi	magːi	moːl	moʔoi
child	*-uaqal	-oqai	-uaka	-uaq	-wake	-oal	-ei
bird	*dVl	dei	daja			‡ul	
bat (n.)	*madel	mədi	mara	merei	madːe	demel*	

Gloss	PAP *l	KLN l	Kui l	ABU Ø	KMN i	SWL l	WRS l
banana	*mogol	mɔgɔl			moːi		mulul
child	*-uaqal	ɔl	ol				ol
bird	*dVl		adol	ruo	atoi	adala	adol
bat (n.)	*madel	mədel	madel	mare	matei		

* With metathesis.

TABLE 39. MEDIAL *r

Gloss	PAP *r	TWA r	NDB l	KER r	WP l	BLG r	ADN l	KLN r	Kui r	ABU j	KMN l	SWL r	WRS r
two	*araqu	raq	raqu	rax-	alaku	aru	aːlu	ərɔk	oruku	ajoku			
water	*jira	jir	jila	ir	‡hila	dʒar	si	arɑ	-eːr	je	ili	iria	ira
sing	*dar(a)	daːr			dali	dar	daːla		dar	yai		dara	dra
bone	—	kir	kili	kiri		ira							
laugh	*jari	-jaɦar			jali	‡ijar	asala		jeri		‡jeːi	jara	jer

14. Some dialects of WP have *r > *l* in both medial and final position, thus WP (Lamma dialect) *bat:al* 'maize'. However, in no dialect of WP is *r preserved as *r*, so forms such as WP *par* 'scorpion' must be borrowings.

Other irregularities in the sets for final *r are due to borrowing. Adang has many forms such as *hu:l* 'moon' with final *l* where *j* is expected as a reflex of *r; these are likely borrowings from neighboring Kabola. In addition to *hu:l* < Kabola *wul,* consider *bɛl* 'dog' < Kabola *bel* and *du:l* 'rat' < Kabola *dul.*

4.6 GLIDES. We reconstruct two glides, *w and *j, in PAP. In most languages, *w is preserved as *w* in all positions. In initial position, only BLG *v* < *w reflects a change; other languages preserve *w, as can be seen in table 41. In ADN *w is phonemically /w/ but may be realized as [f] or [w].

We find insufficient evidence to reconstruct *w in noninitial position. Potential correspondences representing noninitial *w are likely either to be underlying vowels or to reflect original initial *w. For example, the stem-intial consonant in the word for 'ear' (see table 40) is usually analyzed as a glide in KLN, Kui, ABU, KMN, SWL, and WRS. However, regardless of the synchronic analyses, these forms are likely to reflect an original vocalic form, and we reconstruct PAP *uar(i). Apparent medial *w is also found in the word for 'lime': KER *awar*, WP *hauwe*, BLG *avar*, ADN *a:wi*, ABU *awai*, KMN *awoi*. This correspondence matches that for initial *w and even supports reconsruction of PAP *hawar 'lime'. However, this form is likely to be an original compound; compare *war 'stone'. Another example of a potential compound containing medial *w is found in the word for 'five' (see table 27).

TABLE 40. FINAL *r

Gloss	PAP *r	TWA r	NDB Ø	KER r	WP Ø	BLG r	ADN i
stone	*war	war	‡wala	war		var	fui
scorpion	*pVr	par		par		per	pai
tail	*-or(a)	-or	‡-ola	-or		ora	olo?
ear	*-uar(i)	-uar	-ow	-uar	-ue	‡-veli	
lime	*hawar	ħor	wa	awar	hauwe	avar	a:wi
maize	—	batar	ba:ta	batar	bat:e	batar	bati?
moon	*wur	wur	hula	ur		wuru	‡hu:l

Gloss	PAP *r	KLN r	Kui r	ABU i	KMN i	SWL r	WRS r
stone	*war	wɔr	wor	vi:	woi	wara	wor
scorpion	*pVr		per	pei	‡fal		per(buk)
tail	*-or(a)	-or	-or		-(w)ui	-(w)o:ra	(w)ori
ear	*-uar(i)	-uɛr	‡-uel	-uei	-uai	uari	-ueri
lime	*hawar	əwɛr	o:r	awai	awoi		or
maize	—		batar	fati	patei	patar	peter
moon	*wur	ʊr	ur		wui		urak

TABLE 41. INITIAL *w

Gloss	PAP *w	TWA w	NDB w	KER w	WP w	BLG v	ADN w	KLN w	Kui w	ABU w	KMN w	SWL w	WRS w
sun	*wad(i)	war (get)	weri	wer		ved	fɛd			wari	wati	wadi	widi
blood	*wai	wai	we	we:	wai	ve	fi:	wɛ?	we	wea	we:	wi:	wei
stone	*war	war	wala	war		var	fui	wɔr	wor	wi	woi	wara	wor
bathe	*weli	wei		-wei-		vela	wili	-wɛl	weli	-wel	-wei	-wile	-weli

Apparent reflexes of final *w are better analyzed as reflecting original vowels, for example, WP *lau*, ADN *lowo?,* ABU *lou*, SWL *lu*, and WRS *aloi* 'bark (v.)'. Without additional supporting evidence we do not reconstruct *w in final position.

The initial reflexes of the palatal glide *j are relatively straightforward once a few simple rules are taken into account. In KER, ADN, Kui, KMN, SWL, and WRS the reflex of *j is lost before a high front vowel [i]. In WP, it becomes *h* in the same environment. This is illustrated in table 42.

ADN *si* 'water' appears to be an exception, since *j is not lost in spite of the following high front vowel. However, other dialects of ADN have *sij*, with a complex vowel sequence that may have initially blocked deletion of *j. In Kui *e:r* 'water', later vowel quality changes have obliterated the environment that triggered loss of *j. NDB and ADN lose the initial syllable of 'dog' because the form had final stress, and in those languages the initial unstressed syllable was lost. WRS *wetiŋ* 'five' irregularly begins with *w* instead of *j*.

We do not reconstruct *j in noninitial position. Where noninitial *j* is found in modern languages we assume this is a reflex of a vowel. One example is NDB *buja* 'betel nut' < *bui.

5. PROTO–ALOR-PANTAR AND SUBGROUPING. Having established regular consonant correspondences, we can now turn to the reconstruction of the PAP consonant system and historical sound changes that give rise to the modern languages. The consonants reconstructed for PAP are given in table 43. The inventory is very similar to that found in many of the modern Alor-Pantar languages, and its size is typical for the East Nusantara region (Hajek 2010). The phonemic status of PAP *r is unclear; hence it is included in parentheses in table 43. As discussed below, *r may have been an allophone of *j.

TABLE 42. INITIAL *j

Gloss	PAP	TWA	NDB	KER	WP	BLG	ADN
	*j	j	j	j	j	dʒ*	s
water	*jira	jir	jila	ir	hila	dʒar	si
bad	*jasi	jas	jeci	jas-	jasa	‡dʒasi	saha†
dog$_1$	—	jiˈvar	bar	iˈbar	jaˈbːe	dʒaˈbar	‡bɛl
five	*jiwesin	jusan	jisin	isin	jasiŋ	‡jisiŋ	iwɛhɛŋ
star	*jibC	jif	jifa	‡ip(alaq)	hibːi		iːb
laugh	*jari	‡jaħar			jali	ijar	asala

Gloss	PAP	KLN	Kui	ABU	KMN	SWL	WRS
	*j	Ø	j	j	j	j	j
water	*jira	arɑ	eːr	je	ili	iria	ira
bad	*jasi					jaːti	
dog$_1$	—						
five	*jiwesin	əwəh	jesan	jetiŋ	iwesiŋ	joːtiŋ	‡wetiŋ
star	*jibC	ib	ib(ra)				
laugh	*jari		jeri		jeːi	jara	jer

* Steinhauer (1995: 268) notes that this is an "alveopalatal stop with weak affrication" for most speakers, while some older speakers pronounce it as a voiced alveolar fricative.

† Denotes 'broken'.

The presence of the uvular stop, though well supported by modern reflexes in several Pantar languages, is highly unusual in the region. Only 2.4 percent of the languages in Maddieson's (2005) survey of consonant inventories contain uvular stops, though two of those languages are Trans-New Guinea (Kunimaipa and Hamtai). This figure is consistent with Hajek's (2010) survey of the phonological systems of 71 languages of East Nusantara. Hajek identifies only one language other than Teiwa that contrasts velar and uvular stops; this is the West Papuan language Tehit. We can discern additional anomalies by examining the distribution of consonants, summarized in table 44.

All consonants except *r may occur in initial position. In contrast, the glides *j and *w do not occur in medial and final positions; final glides in the modern languages derive from original vowels. The complementary distribution of *r and *j raises the possibility that *r was actually an allophone of *j in PAP. This would make PAP more typically Papuan, in that Papuan languages are said to usually lack an /r/ : /l/ distinction (cf. Foley 1986). We indicate this possibility by listing *r* in parentheses in tables 43 and 44. It should also be noted that while *g does occur in initial position, it occurs there only in pronominal forms.

The absence of stops *g, *p, and *q in final position is striking. As noted in 4.1 and 4.2, we actually do find some marginal examples of reflexes of final *g and *q in some of the languages—for example, TWA *lia:g* 'rattan' and KER *banax* 'smoke'—so it may well be that the lack of evidence for final *g and *q is an artifact of our limited data set. However, the lack of final *p is robustly evidenced in our data. All instances of final *p* in the modern languages can be traced to either an original medial *p, as in TWA *tap* < *tapai 'pierce', or an original voiced stop *b, as in WP *hap* < *hab(i) 'fish'.

The reconstruction of this set of protophonemes implies that certain sound changes have occurred in the daughter languages. Examining the correspondence sets in the previous section we identify seventeen sound changes that are each shared by at least two languages. As is readily apparent from the distribution of the changes shown in table 45, many of these changes are cross-linguistically common, and hence may be of marginal

TABLE 43. RECONSTRUCTED PAP CONSONANT INVENTORY

	LABIAL	ALVEOLAR	PALATAL	VELAR	UVULAR	GLOTTAL
STOP	p b	t d		k g	q	
FRICATIVE		s				h
NASAL	m	n				
GLIDE	w		j			
LIQUID		l (r)				

TABLE 44. DISTRIBUTIONAL RESTRICTIONS ON PAP CONSONANTS

	INITIAL	MEDIAL	FINAL		INITIAL	MEDIAL	FINAL
b	+	+	+	s	+	+	+
d	+	+	+	h	+	–	–
g	+	+	–	m	+	+	+
p	+	+	–	n	+	+	+
t	+	+	+	l	+	+	+
k	+	+	+	(r)	–	+	+
q	+	+	–	j	+	–	–
				w	+	–	–

value for subgrouping, for they may have occurred independently. Additionally, many of the changes cross-cut each other, further complicating internal subgrouping. For example, the change *s > *h* groups ADN with BLG and KLN, while the change *r > *l* groups ADN with NDB, WP, and ABU.

The most widespread of these changes is *h > Ø, which occurs in all languages except TWA and WP. However, this change is typologically common and may have occurred independently in several languages. We choose not to base subgrouping on this change. The second most widespread of these changes is *q > k, which occurs in all languages except the Pantar languages TWA, NDB, and KER. This change results in a merger of *k and *q in most daughter languages, while TWA, NDB, and KER keep these phonemes distinct. However, closer examination reveals that WP also distinguishes reflexes of *k and *q, though not in all positions. WP, as noted previously, geminates original stops in medial position, with the exception of *q. Thus, in medial position *k > WP *k:* and *q > WP *k*. Using this evidence to support WP as maintaining the distinction between *k and *q, we can then identify a large group of languages that merge these phonemes. The eight languages so identified are precisely the languages of Alor and the Pantar Strait, namely, BLG, ADN, KLN, Kui, ABU, KMN, SWL, and WRS. We take this change to define an Alor subgroup.

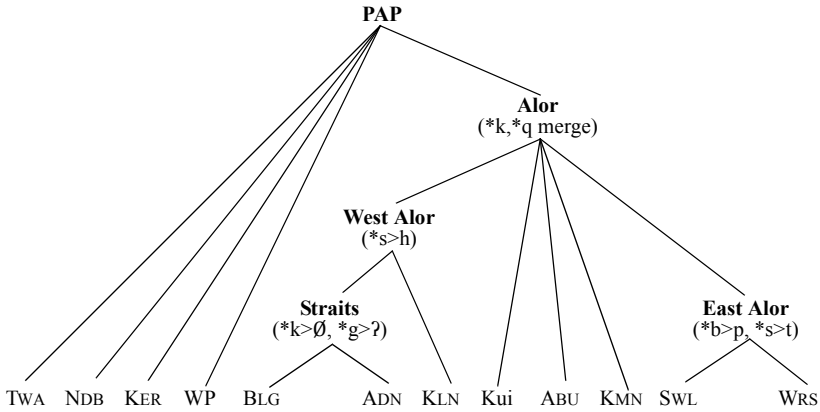
Within the Alor group we can distinguish two lower level subgroups. In the east, the languages SWL and WRS share the innovations *b > *p* and *s > *t*. The former change is also shared with KMN, the latter with ABU. So while it is tempting to expand this group, only SWL and WRS share both of these innovations, defining a subgroup we refer to as East Alor. In the west, the languages BLG and ADN share innovations *k > Ø, *g > *ʔ*, and *s > *h*, defining a group we label Straits. The latter change is also shared with KLN, providing weak support for an intermediate grouping that we label West Alor. The remaining changes cross-cut these and do not provide additional subgrouping information. See figure 2.

The tree based on shared phonological innovations differs in several ways from previous classifications based on lexicostatistics. In particular, while the eastern languages

TABLE 45. SOUND CHANGES FOUND IN AT LEAST TWO LANGUAGES

CHANGE	LANGUAGES
*b > f	TWA, NDB, ABU (in TWA and NDB only noninitially)
*b > p	KMN, SWL, WRS
*d > r	ABU, Kui (in Kui only finally)
*g > ʔ	BLG, ADN
*k > Ø / _#	BLG, ADN
*q > k	WP, BLG, ADN, KLN, Kui, ABU, KMN, SWL, WRS (ADN ʔ < k < *q)
*s > h	BLG, ADN, KLN
*s > t	ABU, SWL, WRS
*h > Ø	everywhere but TWA and WP
*m > ŋ / _#	WP, BLG, ADN
*n > ŋ / _#	NDB, KER, WP, BLG, ADN, ABU, KMN SWL, WRS
*l > i / _#	TWA, KER, ADN, KMN
*l > Ø / _#	NDB, WP, ABU
*r > l / V_V	NDB, WP, ADN, KMN
*r > Ø / _#	TWA, KER, WP
*r > i / _#	BLG, Kui, ABU

FIGURE 2. SUBGROUPING OF ALOR-PANTAR BASED ON SHARED PHONOLOGICAL INNOVATIONS



SWL and WRS form a subgroup, they do not constitute primary branches from PAP, as has been suggested in several previous classifications (cf. Wurm 1982, Lewis 2009).

Having reconstructed the consonant system, we can proceed with a reconstruction of PAP vocabulary. Although we identify 109 distinct lexical correspondences in our data set, not all correspondences are widely attested across the full range of languages. We reconstruct vocabulary items only when reflexes can be found in at least one language of Pantar (TWA, NDB, KER, WP), one language of western Alor and the Pantar Strait (BLG, ADN, KLN, Kui), and one language of eastern Alor (ABU, KMN, SWL, WRS). We exclude from reconstruction very obvious recent borrowings, such as 'maize', but we include some forms that are older Austronesian (AN) borrowings, such as 'pig', 'betel nut', and 'betel vine'. We know that these items/animals were introductions that roughly coincide with the arrival of the AN languages in the area. The fact that these loans can be reconstructed and show regular sound correspondences can be taken as evidence for the claim that the split-up of PAP followed the arrival of AN in the region (perhaps 4,500–4,000 BP; Bellwood 1997:123, Pawley 1998:684–85, Ross 2005:42). However, it is equally likely for later diffusions to exhibit patterns very much like regular sound correspondences. Settling this matter requires independent evidence dating PAP relative to AN.

Table 46 lists 97 vocabulary items that can be reconstructed at the level of PAP on this basis. Since the focus of our reconstruction is on the consonants, the vowels in the reconstructed vocabulary should be interpreted with caution. We do not make any strong claim regarding the nature of the PAP vowel system.

6. WIDER GENETIC AFFILIATIONS. The reconstruction of PAP puts us in a much better position to assess the external relationships of the family. Rather than comparing individual Alor-Pantar languages without knowledge of the language-particular etymologies, we can now begin to compare PAP forms with other languages and reconstructed protolanguages. A full study of these external relationships is beyond the scope of this paper, but we consider some possibilities for future investigations here.

Comparison with Papuan languages of Timor has until recently been dogged by lack of sufficient data on both AP and Timor languages. While documentation of all the Timor languages is now underway, identifying cognates remains difficult due to significant relexification, motivated in particular by contact with Austronesians. Items borrowed from Austronesian languages dominate basic semantic domains such as kin, governance, material culture, and agriculture across the Papuan languages of Timor (see, for example, Huber 2011:40–43, McWilliam 2007, Schapper 2010:22–25). For instance, on a basic Swadesh 200-word list for Bunaq, 40 items can be identified as borrowed from neighboring Austronesian languages. Nevertheless there do appear to be cognates linking AP and Timor languages. As yet, there is no clear primary subgrouping emerging from sound changes in these data, but the weight of lexical and morphological evidence points to Bunaq being more closely linked to the eastern Timor languages than to the AP languages (cf. Schapper 2011). However, the observed similarities may simply represent innovations in PAP and features of PTAP conserved in Timor, rather than subgroup-

TABLE 46. RECONSTRUCTED PAP VOCABULARY

*-ain(i,u)	‘name’	*jari	‘laugh’	*pVr	‘scorpion’
*aman	‘thatch’	*jasi	‘bad’	*qaba(k)	‘spear’
*aqana	‘black’	*jibC	‘star’	*qar-	‘tens’
*-ar	‘vagina’	*jira	‘water’	*qin	‘mosquito’
*araqu	‘two’	*jira(n)	‘fly’ (v.)	*siba	‘new’
*asi	‘bite’	*jiwesin	‘five’	*talam	‘six’
*bagai	‘crocodile’	*kusin	‘fingernail’	*tam	‘saltwater’
*balin	‘axe’	*kVt	‘flea’	*tama	‘fat’
*baj	‘pig’	*-leb(ur)	‘tongue’	*-tan	‘hand/arm’
*bis	‘mat’	*luk(V)	‘crouch’	*tapai	‘pierce’
*bob	‘wave’	*lVsi	‘monitor lizard’	*tas	‘stand’
*bui	‘betel nut’	*madel	‘bat’ (n.)	*tei	‘tree’
*bukan	‘guard’	*mai	‘come’	*temVk	‘bedbug’
*bunaq	‘smoke’	*mait	‘betel vine’	*ten	‘ripe’
*dar(a)	‘sing’	*mari	‘bamboo’	*tia	‘recline’
*dul(a)	‘slippery’	*mi	‘(be) in/on’	*tiara	‘expel’
*dumV	‘thick’	*mid	‘climb’	*-tiari(n)	‘close’ (v.)
*dur	‘rat’	*-mim	‘nose’	*-tok	‘stomach’
*dVl	‘bird’	*minV	‘die’	*tukV	‘short’
*-ena	‘give (to s.o.)’	*mis	‘sit’	*u:b	‘sugarcane’
*ga-	3SG	*mogol	‘banana’	*-uaqal	‘child’
*ge-	3GEN	*mudi	‘body hair’	*-uar(i)	‘ear’
*gi-	3PL	*mudin	‘plant’ (v.)	*-uas	‘teeth’
*ha-	2SG	*-muk	‘horn’	*uku	‘knee’
*hab(i)	‘fish’	*na-	1SG	*Vde	‘burn’
*haban	‘village’	*nai	‘eat/drink’	*wad(i)	‘sun’
*had(a)	‘fire’	*naN(a)	‘sibling’	*wai	‘blood’
*hami	‘breast’	*nuk	‘one’	*war	‘stone’
*has	‘excrement’	*-od	‘throw’	*wat(a)	‘coconut’
*hasak	‘empty’	*-or(a)	‘tail’	*weli	‘bathe’
*hawar	‘lime’	*p(i,u)nV	‘hold’	*wur	‘moon’
*hipar	‘dream’	*pi-	1pl.incl		
*is(i)	‘fruit’	*purVN	‘spit’		

defining shared innovations in Timor. More data remain to be gathered in specialized domains—for example, botanical and ornithological—as these appear to be more conservative in both AP and Timor languages, before a clearer picture of the genealogy of the TAP languages can be gained.

We can also make use of reconstructed PAP vocabulary to begin to assess the wider genetic relationships beyond neighboring Timor. As noted in section 2, beginning with Wurm, Voorhoeve, and McElhanon (1975), most authors have assumed a genetic affiliation with Trans-New Guinea languages. Lewis (2009) places the Alor-Pantar languages within a West Timor-Alor-Pantar subgroup of the West Trans-New Guinea family. Within this putative family, the West Timor-Alor-Pantar group (our Alor-Pantar plus Bunaq and Adabe) is asserted to be coordinate with the following families: Wissel Lakes, Dani, West Bomberai, and East Timor. Ideally, we would procede to examine the classification of Alor-Pantar by evaluating lexical correspondences between reconstructed protolanguages for each of these subgroups. That is, we would apply the comparative method at the level of the subgroups. This task is complicated by the lack of adequate bottom-up reconstructions for the relevant groups. However, it is possible to compare our reconstructions directly with forms already reconstructed for the wider Trans-New Guinea family. Strictly speaking, this should not be taken to represent an application of the comparative method, since rather than comparing two (or more) intermediate protolanguages for putative subgroups we are instead comparing a single intermediate protolanguage with a putative parent. Nevertheless, given the long-standing assumption that Alor-Pantar languages belong to Trans-New Guinea, we believe even a cursory comparison is justified at this point.

Relatively few PTNG reconstructions have been published. However, Pawley (n.d.) lists 98 reconstructions, each well supported by correspondences having reflexes in at least two of the generally recognized major TNG subgoups. This is by far the most extensive source of PTNG vocabulary available. It represents a top-down reconstruction, and Pawley cautions that he is "not greatly concerned with the precise phonetic realizations" of the obstruents. Hence, it seems fair to allow some phonetic latitude in assessing potential lexical correspondences between PAP and PTNG forms. Table 47 below compares PAP vocabulary with semantically equivalent PTNG reconstructions gleaned from Pawley (n.d.). For this comparison, we draw from the entire set of reconstructed PAP vocabulary in table 46. The forms listed in this table do not reflect an assertion of cognacy or similarity.

Among these 34 forms there are some intriguing similarities, such as PAP *nai, PTNG *na- 'eat'; PAP *hami, PTNG *amu 'breast'; PAP *mai, PTNG *me 'come'; and PAP *tukV, PTNG *tukumba[C] 'short'. If we stretch the phonetic leeway a bit further we might also include 'die', 'excrement', 'nose', and 'tongue' as reflecting similarities. Even the more restricted set represents four out of 34 forms, a sizeable percentage (12 percent) of this (arbitrarily selected) list of basic vocabulary. Nevertheless, while intriguing, this set of forms is too small to support any conclusions about regular phonological correspondences.

Now, it might be argued that the time depth being considered would obscure any potential regular phonological correspondences. But this argument runs counter to available evidence. The spread of TNG has been linked to the spread of agriculture (Bellwood 2001), but agriculture emerges only ca. 10,000 BP in the eastern highlands of New

Guinea (Denham et al. 2003), with a westward spread somewhat later, perhaps around 6,000 BP (Pawley 1998). Such a figure is at the upper limits of the time depth usually expected of the comparative method, but crucially this figure does not place Alor-Pantar outside the bounds of the comparative method. There may well be other types of evidence—such as pronouns or typology—that support a connection between Alor-Pantar and Trans-New Guinea (pace Ross 2005). However, now that good lexical data are available, we can no longer disregard the lexical evidence, and the lexical data do not support a connection between Alor-Pantar and Trans-New Guinea languages.

Crucially, this point should not be taken as a claim that the Alor-Pantar languages do not belong to the Trans-New Guinea family: until an alternate genealogical link is established, such a negative claim will remain impossible to prove. Rather, our claim is a more narrow one: in the absence of supporting evidence, Alor-Pantar should be considered a distinct family unrelated to Trans-New Guinea.

TABLE 47. PAP COMPARED WITH PTNG (Pawley n.d.)

PAP	PTNG	
*balin	*tu	‘axe’
*dVl	*n[e]i, *jaka, *nVma	‘bird’
*wai	*ke(nj,s)a	‘blood’
*hami	*amu	‘breast’
*mai	*me-	‘come’
*minV	*kumV-	‘die’
*-uar(i)	*ka(nz,t)(i,e)[C], *tVmV[d]	‘ear’
*nai	*na-	‘eat/drink’
*has	*ata	‘excrement’
*kusin	*mbutuC	‘fingernail’
*had(a)	*kend(o,u)p, *inda	‘fire’
*jira(n)	*pululu-	‘fly’ (v.)
*-ena	*mV	‘give (to s.o.)’
*-tan	*sa(ᵃg,k)al	‘hand/arm’
*uku	*(ŋg,k)atuk	‘knee’
*jari	*ᵃgiti	‘laugh’
*wur	*takVn[V], *kal(a,i)m	‘moon’
*qin	*kasin	‘mosquito’
*-ain(i,u)	*imbi, *wani	‘name’
*siba	*kVtak	‘new’
*-mim	*mundu	‘nose’
*tukV	*[ka]tumba[C], *tukumba[C], *kumb(a,u)	‘short’
*mis	*məna-	‘sit’
*bunaq	*kambu(s,t)(a,u), *(kambu-)la(ᵃg,k)a	‘smoke’
*tas	*t(a,e,i)k[V]-	‘stand’
*war	*ka(mb,m)u[CV],*[na]muna	‘stone’
*wad(i)	*kamali, *ketane	‘sun’
*-or(a)	*a(mb,m)u	‘tail’
*-uas	*maᵃgat[a], *(s,t)i(s,t)i	‘teeth’
*-leb(ur)	*mbilaᵃ, *me(l,n)e	‘tongue’
*tei	*inda	‘tree’
*araqu	*ta(l,t)(a,e)	‘two’
*jira	*ok[V], *nVk, nVL	‘water’
*ha-	*ᵃga[k]	2SG

7. OUTLOOK. The examination of sound correspondences across the Papuan languages of Alor and Pantar supports the view that these languages are genetically related. Our application of the comparative method confirms the conclusions of Stokhof's (1975) lexicostatistical analysis and provides a solid foundation from which to advance future research on the linguistic prehistory of the region. The Alor-Pantar language family is a Papuan outlier, surrounded by Austronesian languages and located some 1000 km. from the Papuan languages of the New Guinea mainland. The question of whether these languages predate the arrival of Austronesian languages in the region remains unanswered. The nature of the relationship with the Papuan languages of Timor invites further study. The reconstructed PAP vocabulary and sound system presented in this paper will be a key tool in this future work, as it allows us to distinguish inheritance from diffusion.

Our comparative work also leads us to propose new subgroups within Alor-Pantar, which will lead to better understanding of the internal history of the family. The subgrouping structure proposed here suggests an origin and original break-up of the family in East Pantar. But the overall linguistic picture remains extremely complex, defying a model based solely on vertical transfer and inheritance. Widespread multilingualism is the norm in the region, and borrowings from neighboring languages—such as WP *bagis* 'whine' from Deing *bagis* 'cry'[15]—are extremely common. Additionally, genetic studies indicate that East Nusantara and the Alor-Pantar region in particular is a melting pot with a long history of genetic admixture (Mona et al. 2009), and it may well be that the same holds for languages. On the other hand, further data, as well as the further reconstruction of PAP, may allow us to suggest more clearly defined subgroups.

Additional lexical data being gathered currently may provide further clues as to the history of language contact, allowing us to better understand borrowing both between AP languages and from surrounding Austronesian languages. Crucial to this effort is the sorting of lexical data into strata to identify just which lexical items have a higher rate of cognacy (indicating inheritance) and which are readily borrowed. Additional data will also allow us to reconstruct culture history in order to understand the relationship between linguistic history and nonlinguistic events such as the development or introduction of agriculture.

More broadly, this work demonstrates that the comparative method can be effectively applied to bottom-up reconstruction of Papuan languages. Crucial to this effort is access to a reliable documentary record, which in the case of the Alor-Pantar languages now exists. By making this record more available and by attempting an initial bottom-reconstruction of PAP, we hope to inspire further research on the linguistic prehistory of the region.

REFERENCES

Anceaux, J. C. 1973. Postscript to F. S. Watuseke "Gegevens over de taal van Pantar." *Bijdragen tot de Taal-, Land- en Volkenkunde* 129:345–46.

Anonymous. 1914. De eilanden Alor en Pantar, Residentie Timor en Onderhoorigheden. *Tijdschrift van het Koninklijk Nederlandsch Aardrijkskundig Genootschap* 31:70–102.

15. Geographically, Deing is between western Pantar and Teiwa. It appears to be closely related to Teiwa.

Baird, Louise. 2005. Doing the split-S in Klon. In *Linguistics in the Netherlands*, ed. by J. Doetjes and J. van der Weijer, 1–12. Amsterdam: John Benjamins.
———. 2008. *A grammar of Klon: A non-Austronesian language of Alor, Indonesia*. Canberra: Pacific Linguistics.
———. 2010. Grammaticalisation of asymmetrical serial verb constructions in Klon. In *Typological and areal analyses: Contributions from East Nusantara*, ed. by Michael Ewing and Marian Klamer, 189–206. Canberra: Pacific Linguistics.
Bellwood, Peter, 1997. *The prehistory of the Indo-Pacific archipelago.* 2nd ed. Honolulu: University of Hawai'i Press.
———. 2001. Early agriculturist population diasporas? Farming, language, and genes. *Annual Review of Anthropology* 30:181–207.
Berthe, Louis 1963. La morpho-syntaxe de Buna' (Timor Central). *L'Homme* 3:106–16.
Capell, Arthur. 1943a. Peoples and languages of Timor. *Oceania* 14(3):191–219.
———. 1943b. Peoples and languages of Timor. *Oceania* 14(4):311–37.
———. 1944. Peoples and languages of Timor. *Oceania* 15(3):19–48.
———. 1965. The Oirata language. *Lingua* 14:360–70.
———. 1975. The West Papuan phylum: General, and Timor and areas further west. In *New Guinea area languages and language study,* vol. I: *Papuan languages and the New Guinea linguistic scene,* ed. by S. A. Wurm, 667–716. Canberra: Pacific Linguistics.
Cowan, H. K. J. 1957. Prospects of a Papuan comparative linguistics. *Bijdragen tot de Taal-, Land- en Volkenkunde* 113:70–91.
de Josselin de Jong, J. P. B. 1937. *Studies in Indonesian culture I: Oirata, a Timorese settlement on Kisar.* Amsterdam, N. V. Noord Nederlandsche Uitgevers Maatschappij.
Denham, T. P., S. G. Haberle, C. Lentfer, T. Fullagar, J. Field, M. Therin, N. Porch, and B. Winsborough. 2003. Origins of agriculture at Kuk Swamp in the Highlands of New Guinea. *Science* 201:189–93.
Donohue, Mark. 1996. Inverse in Tanglapui. *Language and Linguistics in Melanesia* 27:101–18.
———. 2008. Bound pronominals in the West Papuan languages. *Morphology and language history: In honor of Harold Koch*, ed. by Claire Bowern, Bethwyn Evans, and Luisa Miceli, 43–58. Amsterdam: John Benjamins.
Du Bois, Cora Alice. 1944. *The people of Alor; A socio-psychological study of an East Indian Island.* Minneapolis and St. Paul: The University of Minnesota.
Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
Fox, J. J. 1988. Historical consequences of changing patterns of livelihood on Timor. In *Contemporary issues in development,* vol. 1, ed. by Deborah Wade-Marshall and Peter Loveday, 259–79. Canberra: Research School of Pacific Studies, The Australian National University.
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.
Hajek, John. 2010. Towards a phonological overview of the vowel and consonant systems of East Nusantara. In *Typological and areal analyses: Contributions from East Nusantara*, ed. by Michael Ewing and Marian Klamer, 25–46. Canberra: Pacific Linguistics.
Holton, Gary. 2004. Report on recent linguistic fieldwork on Pantar Island, Eastern Indonesia. Report to the National Endowment for Humanities. Washington, DC.
———. 2008. The rise and fall of semantic alignment in North Halmahera, Indonesia. In *The typology of semantic alignment*, ed. by Mark Donohue and Søren Wichmann, 252–76. Oxford: Oxford University Press.
———. 2010. Person-marking, verb classes, and the notion of grammatical alignment in Western Pantar (Lamma). In *Typological and areal analyses: Contributions from*

East Nusantara, ed. by Michael Ewing and Marian Klamer, 101–21. Canberra: Pacific Linguistics.

Holton, Gary, and Mahalalel Lamma Koly. 2008. *Kamus pengantar Bahasa Pantar Barat (Companion dictionary of Western Pantar)*. Kupang, Indonesia: UBB-GMIT.

Huber, Juliette. 2011. *A grammar of Makalero: A Papuan language of East Timor*. Leiden: University of Leiden.

Klamer, Marian. 2008. The semantics of semantic alignment in Eastern Indonesia. In *The typology of semantic alignment*, ed. by Mark Donohue and Søren Wichmann, 221–51. Oxford: Oxford University Press.

———. 2010a. *A grammar of Teiwa*. Berlin: Mouton de Gruyter.

———. 2010b. One item, many faces: 'Come' in Teiwa and Kaera. In *Typological and areal analyses: Contributions from East Nusantara*, ed. by Michael Ewing and Marian Klamer, 205–27. Canberra: Pacific Linguistics.

———. 2010c. Ditransitives in Teiwa. In *Studies in ditransitive constructions*, ed. by Andrej Malchukov, Martin Haspelmath, and Bernard Comrie, 427–55. Berlin: Mouton de Gruyter.

———. 2011. *A short grammar of Alorese (Austronesian)*. Munich: Lincom.

———. 2012. Reality status in Teiwa (Papuan). *Language Sciences*. Online: http://dx.doi.org/10.1016/j.langsci.2011.08.006.

———. Forthcoming:a. *Kosa kata Bahasa Teiwa–Indonesia–Inggris* (*Teiwa–Indonesian–English glossary)*. Kupang: Language and Culture Unit UBB.

———. Forthcoming:b. Papuan-Austronesian language contact: Alorese from an areal perspective. *Language Documentation and Conservation*, http://nflrc.hawaii.edu/ldc/.

Klamer, Marian, and František Kratochvíl. 2006. The role of animacy in Teiwa and Abui (Papuan). *Proceedings of BLS 32*. Berkeley: Berkeley Linguistic Society.

———. 2010. Abui tripartite verbs: Exploring the limits of compositionality. In *Rara and Rarissima: Documenting the fringes of linguistic diversity*, ed. by Jan Wohlgemuth and Michael Cysouw, 185–210. Berlin: Mouton de Gruyter.

Klamer, Marian, and Antoinette Schapper. Forthcoming. The development of 'give' constructions in the Papuan languages of Timor-Alor-Pantar. *Linguistic Discovery*, http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/issue.

Kratochvíl, František. 2007. *A grammar of Abui: A Papuan language of Alor*. Utrecht: LOT Publications.

Kratochvíl, František, and Benidiktus Delpada. 2008. *Kamus pengantar Bahasa Abui. Abui–Indonesian–English dictionary*. Kupang, Indonesia: UBB-GMIT.

Kratochvíl, František, Sebastian Fedden, Gary Holton, Marian Klamer, Laura C. Robinson, and Antoinette Schapper. 2011. Pronominal systems in Alor-Pantar languages: Development and diachronic stability. Paper presented at the workshop "Person Forms across Time and Space: Divergence or Convergence," 20th International Conference on Historical Linguistics, Osaka, Japan, July 28.

Lewis, M. Paul. 2009. *Ethnologue: Languages of the world*, 16th ed. Dallas: SIL International.

Lichtenberk, Frantisek. 1998. Did speakers of Proto Oceanic chew betel? *Journal of the Polynesian Society* 107:335–63.

Maddieson, Ian. 2005. Consonant inventories. In *The world atlas of language structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie. Oxford: Oxford University Press.

McWilliam, Andrew. 2007. Austronesians in linguistic disguise: Fataluku cultural fusion in East Timor. *Journal of Southeast Asian Studies* 38(2):355–75.

Mona, Stefano, Katharina E. Grunz, Silke Brauer, Brigitte Pakendorf, Loredana Castri, Herawati Sudoyo, Sangkot Marzuki, Robert H. Barnes, Jörg

Schmidtke, Mark Stoneking, and Manfred Kayser. 2009. Genetic admixture history of Eastern Indonesia as revealed by y-chromosome and mitochondrial DNA analysis. *Molecular Biology and Evolution* 26(8):1865–77.

Nicolspeyer, Martha Margaretha. 1940. *De sociale structuur van een Aloreesche bevolkingsgroep*. Rijswijk: Kramers.

Pampus, Karl-Heinz. 2001. *Mué Moten Koda Kĭwan [Kamus Bahasa Lamaholot Dialek Lewolema, Flores Timur]*. Frankfurt: Frobenius Institut.

Pawley, Andrew. 1998. The Trans New Guinea Phylum hypothesis: A reassessment. In *Perspectives on the Bird's Head of Irian Jaya, Indonesia*, ed. by Jelle Miedema, Cecilia Od , and Rien A. C. Dam, 655–90. Amsterdam: Rodopi.

———. 2001. The Proto Trans New Guinea obstruents: Arguments from top–down reconstruction. In *The boy from Bundaberg: Studies in Melanesian linguistics in honour of Tom Dutton*, ed. by Andrew Pawley, Malcolm Ross, and Darrell Tryon, 261–300. Canberra: Pacific Linguistics.

———. 2005. The chequered career of the Trans New Guinea hypothesis: Recent research and its implications. In *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*, ed. by Andrew Pawley, Robert Attenborough, Jack Golson, and Robin Hide, 67–108. Canberra: Pacific Linguistics.

———. n.d. Some Trans New Guinea Phylum cognate sets. MS.

Robinson, Laura C. 2011. The development of religious dialects in Blagar. Paper presented at the 20th International Conference on Historical Linguistics, Osaka, Japan, July 29.

Rodemeier, Susanne. 2003. Tutu Kadire – Erzählen und Erinnern lokalgeschichtlicher Mythen am Tanjung Muna auf Pantar in Ostindonesien. PhD diss., University of Leipzig.

Ross, Malcolm. 2005. Pronouns as a preliminary diagnostics for grouping Papuan languages. In *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*, ed. by Andrew Pawley, Robert Attenborough, Jack Golson, and Robin Hide, 15–65. Canberra: Pacific Linguistics.

Schapper, Antoinette. 2010. Bunaq: A Papuan language of central Timor. PhD thesis, Australian National University.

———. 2011. Finding Bunaq: The homeland and expansion of the Bunaq in central Timor. In *Life and land in Timor: Ethnographic papers*, ed. by Andrew McWilliam and Elizabeth G. Traube, 163–86. Canberra: ANU E-Press.

Schapper, Antoinette, and Marian Klamer. 2011. Plural words in Papuan languages of Alor-Pantar. Paper presented at the Conference on Language Documentation and Linguistic Theory (LDLT) 3, London, November.

Schapper, Antoinette, and Martin Manimau. 2011. *Kamus Pengantar Bahasa Kamang*. Kupang, Indonesia: UBB-GMIT.

Steinhauer, Hein. 1991. Demonstratives in the Blagar language of Dolap (Pura, Alor, Indonesia). In *Papers in Papuan linguistics,* ed. by Tom Dutton, 177–221. Canberra: Pacific Linguistics.

———. 1995. Two varieties of the Blagar language (Alor, Indonesia). In *Tales from a concave world: Liber amicorum Bert Voorhoeve*, ed. by Connie Baak, Mary Bakker, and Dick van der Meij, 269–96. Leiden: Projects Division, Department of Languages and Cultures of South-East Asia and Oceania.

Stokhof, W. A. L. 1975. *Preliminary notes on the Alor and Pantar languages (East Indonesia)*. Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

———. 1977. *Woisika I: An ethnographic introduction.* Canberra: Pacific Linguistics.

———. 1979. *Woisika II: Phonemics.* Canberra: Pacific Linguistics.

———. 1982. *Woisika riddles.* Canberra: Pacific Linguistics.

———. 1984. Annotations to a text in the Abui language (Alor). *Bijdragen tot de Taal-, Land- en Volkenkunde* 140:106–62.
Vatter, Ernst. 1932. *Ata Kiwan. Ungekannte Bergvölker im tropischen Holland: ein Reisebericht.* Leipzig: Bibliographisches Institut A. G.
Watuseke, F. S. 1973. Gegevens over de taal van Pantar: Een Irian taal. *Bijdragen tot de Koninklijk Instituut voor Taal-, Land- en Volkenkunde* 129:340–46.
Wurm, S. A. 1982. *The Papuan languages of Oceania*. Tübingen: Narr.
Wurm, S. A., C. L. Voorhoeve, and Kenneth A. McElhanon. 1975. The Trans New Guinea Phylum in general. In *New Guinea area languages and language study,* vol. I: *Papuan languages and the New Guinea linguistic scene*, ed. by S. A. Wurm, 299–322. Canberra: Pacific Linguistics.
Yen, Douglas E. 1977. Hoabinhian horticulture? The evidence and the questions from northwest Thailand. In *Sunda and Sahul: Prehistoric studies in Southeast Asia, Melanesia and Australia*, ed. by Jim Allen, Jack Golson, and Rhys Jones, 567–99. London: Academic Press.

gmholton@alaska.edu
M.A.F.Klamer@hum.leidenuniv.nl
fkratochvil@ntu.edu.sg
lcrobinson@alaska.edu
a.c.schapper@hum.leidenuniv.nl