

## Chapter 2

# The internal history of the Alor-Pantar language family

Gary Holton

Laura C. Robinson

This chapter demonstrates that the languages of Alor and Pantar share a common origin by applying the comparative method to primary lexical data from twelve languages sampled across the islands of the Alor-Pantar archipelago. More than one hundred proto-Alor-Pantar lexical items are reconstructed. An internal subgrouping based on shared phonological innovations is proposed and is compared to that derived using computational phylogenetic methods. It is argued that the Alor-Pantar group originally came from the region of the Pantar Strait.

### 1 Introduction

In this chapter we review the reconstruction of proto-Alor-Pantar (pAP) based on the comparative method. We then examine the internal relationships of the Alor-Pantar family and discuss several approaches to subgrouping. In the literature, the Alor-Pantar languages are usually considered to belong to the larger Timor-Alor-Pantar (TAP) group, which includes the Papuan languages of neighboring Timor and Kisar. The Alor-Pantar languages form a well-defined subgroup within TAP and share a common history independent of the Timor languages. The relationship between Alor-Pantar and the Timor languages is discussed in the following chapter; the wider historical relationships with languages beyond Timor-Alor-Pantar are discussed in Chapter 4.

The AP languages form one of only two large pockets of non-Austronesian languages in East Nusantara outside New Guinea (the other being North Halmahera, to which AP languages are not related – see Chapter 4). In contrast to



neighboring Timor, all but one of the two dozen or so indigenous languages of the Alor-Pantar archipelago are non-Austronesian.<sup>1</sup> The single Austronesian language, Alorese, clearly has a more recent origin and today occupies only a few coastal outposts in the archipelago (Klamer 2011; 2012). Early reports noted a clear cultural distinction between the “non-indigenous” coastal Alorese speakers and the “indigenous” mountain populations of Alor and Pantar (Anonymous 1914: 75-8). The non-Austronesian character of the languages (as opposed to the cultures) was first recognized for Oirata, a language spoken on Kisar Island, just east of Timor de Josselin de Jong (1937), and shortly thereafter a connection was made to Abui, a language of Alor (Nicolspeyer 1940). The first evidence for the genealogical unity of the AP languages is found in Stokhof (1975), who compiled 117-item comparative wordlists for 34 language varieties. Stokhof proposed a preliminary lexicostatistical classification based on similarity judgments applied to these lexical data, emphasizing that this classification should be considered “very preliminary” (1975: 13). Holton et al. (2012) employed a much larger lexical dataset to identify regular sound correspondences and establish a reconstruction of pAP using the comparative method.

While the results of the comparative method definitively show that the AP languages form a genealogical unit, the identified phonological innovations are typologically common and do not delineate neat subgroups. After reviewing the subgrouping implications of the phonological innovations, we apply computational methods to the same data and are on this basis able to identify internal groupings. Crucially, the lexical data are coded for cognacy based on identified phonological innovations. The resulting tree of AP languages is consistent with an historical scenario whereby AP languages originate in the Pantar Strait. We begin in the following section by reviewing the recent reconstruction of pAP.<sup>2</sup>

---

<sup>1</sup> We exclude here the Austronesian language Sama-Bajo, which is spoken by recent migrants in a single community on the coast of northern Pantar.

<sup>2</sup> The data and reconstructions in §2 and in the Appendix largely follow Holton et al. (2012), except in the following ways. First, we have followed Robinson & Holton (2012) in correcting some minor typos. Some of these corrections result in changes to our analyses as well. For example, retranscription of Teiwa *ki?in* ‘mosquito’ leads us to reconstruct \*kin rather than \*qin. Second, while the data in Holton et al. (2012) are primarily phonetic, here we have tried to use phonological forms where they are known. Third, we have used different dialects of Adang and Blagar (Pitungbang and Dolabang dialects, respectively) to more closely match existing publications for those languages (e.g., the sketch grammars in Schapper 2014). Fourth, we have consulted new evidence from Timor languages (see Chapter 3) to add new reconstructions or to update the pAP reconstructions based on external evidence. That new evidence has caused us to question the reconstructability of \*b and \*d in final position, as discussed below. Finally, we have removed pAP reconstructions for ‘axe’ and ‘comb’ because we now have evidence that these may be Austronesian loanwords that postdate the breakup of pAP.

## 2 Sound correspondences and reconstruction

The surge in documentary field work over the past decade (see Chapter 1) has provided a robust lexical dataset on which to base the reconstruction of pAP. The primary data source used in this chapter is a set of 400-item vocabulary lists collected for twelve different language varieties with broad geographic representation across the archipelago: Teiwa (Tw), Nedebang (Nd), Kaera (Ke), Western Pantar (WP), Blagar (Bl), Adang (Ad), Klön (Kl), Kui (Ki), Abui (Ab), Kamang (Km), Sawila (Sw), and Wersing (We). (See Figure 2 in Chapter 1.) The vocabulary list was tailored specifically to this task, taking into account three specific goals. First, the list contains basic vocabulary such as that found on a Swadesh list, tailored to include items relevant to the East Nusantara cultural and ecological region. Second, the list includes some non-basic vocabulary which may be diagnostic of shared cultural traits. For example, it includes a number of terms relating to agriculture. Our ability or inability to reconstruct these terms provides insight into the culture history of the pAP speakers and thus sheds light on AP prehistory and migration. Third, the list includes items motivated by a need to find further examples of specific sound correspondences, such as ‘village’ and ‘crocodile’, which both contain pharyngeal fricatives in Teiwa. These lists were supplemented by data from published sources and from ongoing field work by members of the EuroBABEL project. 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 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, ‘-’.

Identification of regular consonant correspondences supports the reconstruction of a pAP inventory containing 14 consonants, as shown in Table 1. Each of these consonant reconstructions is supported by correspondence sets for each position (initial, medial, and final) in which that consonant occurs.

While we can identify regular correspondence sets supporting \*r, this phoneme occurs in complementary distribution with \*j. In fact, \*r is the only consonant which does not occur in initial position (see Table 3 below). In contrast, glides \*j and \*w occur only in initial position; final glides in the modern languages derive from original vowels. The complementary distribution of \*r and \*j raises the possibility that \*r is actually an allophone of \*j in pAP. However, the /r/ ~

Table 1: Reconstructed pAP consonant inventory

|           | Labial | Apical | Palatal | Velar | Uvular | Glottal |
|-----------|--------|--------|---------|-------|--------|---------|
| Stop      | p b    | t d    |         | k g   | q      |         |
| Fricative |        | s      |         |       |        | h       |
| Nasal     | m      | n      |         |       |        |         |
| Glide     | w      |        | j       |       |        |         |
| Liquid    |        | l (r)  |         |       |        |         |

/l/ distinction is found in all of the modern AP languages (see Table 2 below), and at least two liquids must be reconstructed to proto-Timor-Alor-Pantar, the immediate parent of pAP (see Chapter 3).

Although a uvular stop is found in only two of the modern AP languages, the reconstruction of pAP \*q is well-supported by a number of correspondence sets. In addition to Teiwa and Nedebang, which maintain \*q as /q/, in Kaera the reflex of \*q is /x/, which is distinct from the reflex of \*k as /k/. Western Pantar collapses \*q and \*k as /k/ in initial position but maintains the distinction in medial position, as only the reflex of \*k is geminate. Outside these four Pantar languages, the \*q/\*k distinction is lost. From a typological perspective the presence of the uvular stop is highly unusual. Only 2.4% 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 which contrasts velar and uvular stops; this is the West Papuan language Tehit.

The inventory of pAP consonants 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). Most modern AP languages differ in having a velar nasal, which is not reconstructed for pAP. As noted above most AP languages also distinguish /r/ and /l/. The consonant inventory for pAP can be compared with that for the modern language Western Pantar in Table 2.

The Western Pantar inventory exhibits several features typical of phonological developments in the modern languages. First, the uvular stop has merged with the velar stop. Second, Western Pantar has developed a velar nasal in final position. Third, Western Pantar has developed a phonemic glottal stop. The distributional restrictions on pAP consonants are summarized in Table 3. It should be noted that while \*g does occur in initial position, it occurs there only in pronominal forms.

2 *The internal history of the Alor-Pantar language family*

Table 2: Western Pantar consonant inventory

|           | Labial | Alveolar | Palatal | Velar | Glottal |
|-----------|--------|----------|---------|-------|---------|
| Stop      | p b    | t d      |         | k g   | ʔ       |
| Fricative |        | s        |         |       | h       |
| Nasal     | m      | n        |         | ŋ     |         |
| Glide     | w      |          | j       |       |         |
| Liquid    |        | l r      |         |       |         |

Table 3: Distributional restrictions on pAP consonants

|     | Initial | Medial | Final |
|-----|---------|--------|-------|
| b   | +       | +      | (+)   |
| d   | +       | +      | (+)   |
| g   | +       | +      | -     |
| p   | +       | +      | -     |
| t   | +       | +      | +     |
| k   | +       | +      | +     |
| q   | +       | +      | -     |
| s   | +       | +      | +     |
| h   | +       | -      | -     |
| m   | +       | +      | +     |
| n   | +       | +      | +     |
| l   | +       | +      | +     |
| (r) | -       | +      | +     |
| j   | +       | -      | -     |
| w   | +       | -      | -     |

In Holton et al. (2012), we reconstruct the voiced stops \*b and \*d for all positions, but after revising the reconstructions based on external data from Timor (see Chapter 3), many reconstructions which formerly ended in a voiced stop have now been revised to include further segments (i.e., ‘fire’, ‘fish’, ‘sugarcane’, ‘sun’, ‘throw’). The voiced stops \*b and \*d are now only weakly attested in final position in our data, and it is possible that with further evidence, those few reconstructions with final \*b and \*d may need to be revised as well. Note that we also do not reconstruct \*g in final position, so pAP may have had a restriction on final voiced stops, though these do occur in modern AP languages (e.g., Teiwa *lia:g* and Kaera *le:g* ‘rattan’).

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 or to \*b, as in Teiwa *tap* < \**tapai* ‘pierce’, or Western Pantar *hap* < \**habi* ‘fish’.

Drawing from a comparative lexical database consisting of approximately 400 items we identify 129 cognate sets reflecting regular sound correspondences (see Appendix). There are only 127 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 below. However, it is important to note that several of the cognate sets cannot be reconstructed to pAP, since they are found only in a geographically restricted area. That is, in some cases lexemes appear to have been innovated. This is particularly obvious for those meanings for which we have two correspondence sets (distinguished below with subscript numerals). The supporting data for each of these sets is provided in the course of demonstrating the correspondences in the following subsections. The complete set of correspondences can be found in the Appendix.

In this section we describe the 35 consonant correspondences which 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. The tables below set out the consonant correspondences, as well as the reconstructed 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. A dash (-) 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.

## 2 The internal history of the Alor-Pantar language family

Transcription follows IPA conventions.<sup>3</sup> Geminate consonants and long vowels are indicated with a length mark ( : ). Word stress is transcribed here only where relevant to the correspondence in question (e.g., ‘dog<sub>1</sub>’). In most of the modern languages stress is on the penultimate syllable; however, stress may also be attracted to heavy syllables, as in Teiwa *ji’var* ‘dog’. In addition, stress may be phonemically contrastive in some languages, as in Western Pantar *ba’wa* ‘conch shell’ vs. *bawa* ‘drum’.

In the tables the languages are arranged in order roughly from west to east with the western-most languages on the left and the eastern-most languages on the right. This arrangement is maintained throughout all the tables in the paper.

In the following subsections 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 which show the expected form in initial position but an unexpected reflex in medial or final position. Nevertheless, such irregular forms are included in correspondence sets for the sake of completeness. In these cases, the irregular forms are denoted with a preceding double dagger ( † ) in the Appendix. 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, we reconstruct pAP forms only when we have broad geographic evidence. That is, reflexes must be found in minimally one language of Pantar (Teiwa, Nedebang, Kaera, Western Pantar), one language of West Alor and the Pantar Strait (Blagar, Adang, Klou, Kui), and one language of East Alor (Abui, Kamang, Sawila, Wersing). Where reflexes are found only in a restricted region such as Pantar or Eastern Alor, we do not reconstruct a pAP lexeme.

### 2.1 Voiced stops

We reconstruct three voiced stops in labial, apical, and velar positions. Labial and apical voiced stops are well attested in initial and medial positions, and only weakly attested in final position. The evidence for a voiced velar stop in initial position is based entirely on third person pronominal forms, and there is no support for a velar stop in final position.

Initial pAP \*b is retained everywhere except Abui, where it weakens to /f/, and the Eastern Alor languages Kamang, Sawila, and Wersing, where it is devoiced as /p/. This correspondence is found in ‘pig’, ‘betel nut’, ‘axe’, ‘maize’, and

---

<sup>3</sup> 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*.

Table 4: Alor-Pantar voiced stop correspondences

| pAP | Env | Tw  | Nd  | Ke | WP | Bl  | Ad | Kl | Ki | Ab | Km | Sw | We |
|-----|-----|-----|-----|----|----|-----|----|----|----|----|----|----|----|
| *b  | #_  | b   | b   | b  | b  | b   | b  | b  | b  | f  | p  | p  | p  |
| *b  | V_V | ɸ/v | f/v | b  | b: | b   | b  | b  | b  | f  | f  | p  | p  |
| *b  | _#  | ɸ/v | f/v | b  | p  | b   | b  | b  | b  | ∅  | 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  | ?   | ?  | g  | g  | h  | g  | g  | g  |
| *g  | V_V | ħ   | x   | g  | g: | ∅/? | ?  | g  | g  | h  | ∅  | j  | l  |

Table 5: Alor-Pantar voiceless stop correspondences

| pAP | Env | Tw | Nd  | Ke | WP | Bl  | Ad | Kl | Ki | Ab | Km | Sw | We |
|-----|-----|----|-----|----|----|-----|----|----|----|----|----|----|----|
| *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   | ∅  | t  | t  | t  | t  | t  | t  |
| *k  | #_  | k  | k   | k  | k  | k   | ?  | k  | k  | k  | k  | k  | k  |
| *k  | V_V | -  | k   | k  | k: | k   | ?  | k  | k  | k  | k  | k  | k  |
| *k  | _#  | k  | k   | k  | k  | ∅   | ∅  | k  | k  | k  | k  | -  | ∅  |
| *q  | #_  | q  | q   | x  | k  | k/? | ?  | k  | k  | k  | k  | k  | k  |
| *q  | V_V | q  | q   | x  | k  | k   | ∅  | k  | k  | k  | k  | k  | k  |

Table 6: Alor-Pantar fricative correspondences

| pAP | Env | Tw  | Nd   | Ke | WP | Bl | Ad | Kl | Ki | Ab | Km | Sw | We |
|-----|-----|-----|------|----|----|----|----|----|----|----|----|----|----|
| *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  | h  | h  | h  | s  | t  | h  | t  | t  |
| *h  | #_  | h/ħ | ∅    | ∅  | h  | ∅  | ∅  | ∅  | ∅  | ∅  | ∅  | ∅  | ∅  |



‘crocodile’. Thus, Abui *fe*, Kamang *pe*, Sawila *pi*, Wersing *pei* < pAP \*baj ‘pig’. While the correspondence sets for initial \*b are extremely regular, they are not without problems, since they may reflect borrowings. The clearest instance of this problem occurs with ‘maize’, which was first introduced into the region by the Dutch in the 15-16th 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 proximate source for ‘maize’ lexemes in AP (e.g., Tetun *batar* ‘maize’). We do not reconstruct a word for ‘maize’ to pAP, but the cognate set is included here because its consonant correspondences follow the established patterns. That is, the phonological innovations affecting pAP initial \*b and final \*r must postdate the introduction of the lexical item to Alor-Pantar.

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 (i.e., ‘betel’ is reconstructable 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 (e.g., Tetun *bua* ‘betel’, Tokodede *buo* ‘betel’) suggests that AP ‘betel’ lexemes may in fact be borrowings from Austronesian. Given this lexical likeness and the uncertainty of the timing of the arrival of betel in the region, we tentatively reconstruct a pAP (loan) lexeme for ‘betel nut’.

Medial reflexes of \*b are found in ‘village’, ‘dog<sub>1</sub>’, ‘spear’, ‘star’, ‘fish’, ‘tongue’, ‘sugarcane’, ‘shark’, ‘leg’, and ‘new’. These follow the same pattern as initial \*b except in Teiwa, Nedebang, Western Pantar, and Abui. In Teiwa and Nedebang \*b weakens to a fricative; thus, Teiwa *haφan*, Nedebang *afan* < pAP \*haban ‘village’. In Western Pantar \*b geminates in medial position, thus, Western Pantar *hab:aŋ* ‘village’. If the final vowel is lost, \*b is reflected as /p/ in Western Pantar and is lost in Abui (e.g., Western Pantar *hap* < pAP \*habi ‘fish’).

Medial gemination is a characteristic feature of Western Pantar; most pAP stops (including nasal stops) are geminated in medial position in Western Pantar (transcribed here as long consonants *bː*, *dː*, etc.). We infer that modern non-geminate medial stops in Western Pantar reflect either borrowing or innovation that took place after the gemination process. In modern Western Pantar there is a robust phonemic contrast between geminate and non-geminate 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 Western Pantar do we find geminates as a regular reflex of pAP medial stops; elsewhere they occur only sporadically.

Evidence for \*b in final position is based only on a single reconstruction for ‘wave’. In Holton et al. (2012), ‘fish’, ‘tongue’, ‘sugarcane’, and ‘wave’ were all reconstructed with a final \*b, but with many of the cognates containing final vowels which we assumed were epenthetic. External evidence from Timor languages (see Chapter 3 and Schapper, Huber & van Engelenhoven 2012) has forced us to reconstruct those final vowels for ‘fish’, ‘tongue’, and ‘sugarcane’. It is possible that ‘wave’ also had a final vowel in pAP, though we have insufficient evidence to reconstruct that at this time.

The variation in Teiwa and Nedebang between voiced and voiceless reflexes of non-initial \*b appears to be unconditioned. Nedebang *bova* ‘wave’ has a voiced fricative, while *a:fi* ‘fish’ has a voiceless fricative. Klamer (2010: 38) notes that while /ɸ/ and /v/ are distinct phonemes in Teiwa, the voiced variant is quite rare. The sporadic voicing seen in these correspondence sets may reflect a recent phonemicization of /v/.

In initial and medial positions \*d is reflected as /r/ in Abui and as /t/ in Kamang. The other languages retain /d/, with the exception of Western Pantar which has a geminate in medial position, as expected. When the final vowel is lost, Teiwa, Nedebang, Western Pantar, and Kui reflect \*d > r.

Initial \*d correspondences are found in ‘rat’, ‘sing’, ‘bird’, and ‘slippery’. Thus, Teiwa *dur*, Western Pantar *di*, Abui *rui*, Kamang *tui* < pAP \*dur ‘rat’. Medial \*d correspondences are found in ‘to plant’, ‘bat’, ‘right (side)’, ‘throw’, ‘fire’, ‘sun’, and ‘body hair’. Thus, Teiwa *mədi*, Western Pantar *mad:e*, Abui *marel*, Kamang *matei* < pAP \*madel ‘bat’. The unexpected Kaera form *wer* ‘sun’ is likely a borrowing from neighboring Teiwa or Nedebang, as \*d is more regularly reflected as /d/ in final position, as in *od* ‘throw’ and *ad* ‘fire’. On the other hand, Nedebang *mara* and Kaera *merai* ‘bat’ unexpectedly have /r/ in medial position. These forms may reflect a borrowing (from Abui); alternatively, these forms may have a more complex history in which final syllables were originally lost, leading these forms to be treated as final.

Evidence for \*d in final position is based on a single correspondence set for ‘garden’, which is not even reconstructed to pAP. As with \*b, evidence for final \*d was considerably weakened in light of external evidence from Timor. The forms ‘throw’, ‘fire’, and ‘sun’ have all been revised to contain final vowels.

## 2 The internal history of the Alor-Pantar language family

Initial \*g is reflected as a glottal stop in Blagar and Adang, and as a glottal fricative in Abui. However, the reconstruction of initial \*g hinges entirely on the correspondence of third person prefixal forms in pAP. These forms exhibit vowel grading which distinguishes singular, plural, genitive, and locative. In particular, all instances of initial /g/ in modern AP languages can be traced to third person pronouns. Only the third-singular bound pronoun \*ga has reflexes in all languages. The third plural is attested in a few languages and can be tentatively reconstructed as \*gi-. It is absent in the modern AP languages Adang, Klon, Kamang, and Abui, 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 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 (Teiwa, Nedebang, Kaera, Western Pantar, Blagar), and we do not reconstruct it to pAP. Note that Kamang has a regular change of initial \*g to /w/ before back vowels, hence the form wo-.

With some possible exceptions, these forms are bound, occurring as prefixes with either nominal or verbal roots. Exceptions include Adang ?e and Klon ge 3GEN.<sup>4</sup> 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).

The reflexes of medial \*g are much more varied, but they are robustly attested in ‘yellow’, ‘yawn’, ‘banana’, ‘garden’, ‘crocodile’, and ‘hear’. Only in Kaera, Klon, and Kui is \*g retained unchanged in medial position. In Western Pantar we find the expected geminate in all forms except *bagai* ‘crocodile’, which may be a borrowing from Kaera. In Teiwa 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 Sawila and Wersing are supported by only one lexical item each.

The evidence for \*g in final position is extremely weak. In the modern languages final g occurs only in the Pantar languages Teiwa and Kaera (as well as Sar, not in our sample). In our 400-item wordlist, final g is found in only eleven distinct Teiwa word forms. None of these has cognates in a central or eastern

---

<sup>4</sup> 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).

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, Teiwa *mia:g*, Nedebang *mia:gi*, Kaera *miag* ‘yesterday’; and Teiwa *bog*, Nedebang *boga*, Western Pantar *bog:a* ‘young’. Hence, it seems plausible to conclude that Teiwa and Kaera final *g* actually derive from medial \**g* and that pAP \**g* was not found in final position.

## 2.2 Voiceless stops

We reconstruct three voiceless stops in labial, apical, and velar positions. While all the modern languages have glottal stops, as in Adang *ʔahap* ‘to cry’ versus *ahap* ‘jungle’, we do not find sufficient evidence at this time to reconstruct glottal stop to pAP. In initial and medial positions pAP \**p* remains unchanged in all the languages except Kamang, where it weakens to /*f*/; Western Pantar, where it predictably geminates in medial position; and Wersing, where evidence from a single correspondence (‘pierce’) suggests that \**p* was lost in medial position. Correspondence sets reflecting \**p* include ‘hold’, 1PL.INCL, ‘scorpion’, ‘pierce’, and ‘search’. Thus, pAP \**p*{*i,u*}nV > Teiwa *pin*, Blagar *pina*, Adang *puin*, Abui *pun*, Kamang *fun*, Sawila *puni* ‘hold’. The devoicing of \**b* in Sawila and Wersing results in merger of \**b* and \**p*. Note that Western Pantar *par* ‘scorpion’ must be a loan from a language which preserves final \**r*. We find no evidence to support reconstruction of \**p* in final position. Rather, final /*p*/ in modern languages results from loss of final vowels (e.g., \**tapai* > Teiwa *tap* ‘pierce’, \**habi* > Western Pantar *hap* ‘fish’).

In initial and medial positions \**t* remains unchanged in all languages, with the exception of Western Pantar, which has a geminate medially as expected. The reconstruction of initial \**t* is supported by correspondence sets for ‘recline’, ‘saltwater’, ‘short’, ‘stand’, ‘ripe’, ‘far’ and ‘tree’. Reconstruction of medial \**t* is supported by correspondence sets for ‘dry’, ‘maize’, ‘hearth’, and ‘hand/arm’. Reflexes of ‘dry’ and ‘hearth’ are not sufficiently widely distributed to justify reconstruction at the level of pAP. The set for ‘dry’ is found only in the Pantar Strait and Central Alor languages, while the form for ‘hearth’ is found only in the Pantar languages. 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 *tanan* ‘hand’ is superficial only and cannot be taken to indicate that the AP lexemes are Austronesian borrowings. The form *tanan* for ‘hand’ is a lexical innovation of Malayic and cannot be reconstructed to higher levels of the Austronesian family: proto-Malayo-Polynesian (and proto-Austronesian) reconstructions are \**lima* for ‘hand’ and \**baRa* for ‘arm’, and it is

reflexes of these proto-lexemes for ‘hand’ and ‘arm’ which 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.<sup>5</sup> As such, we unproblematically reconstruct pAP \*tan ‘hand/arm’.

Final \*t is preserved in all languages except Adang, where it is lost. The reconstruction of final \*t is supported by cognate sets for ‘leg’, ‘flea’, ‘betel vine’, and ‘wound’. However, only in Teiwa, Kaera, Blagar, and Kui does the reflex of \*t still occur finally. This leads to some uncertainty as to whether these forms may have been originally medial. The correspondence set for ‘betel vine’, for example, as it is reflected as 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 Western Pantar 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 Adang, \*t > h /Vi\_ in Klón (presumably via *s*), and \*t > s/ Vi\_ in Kui, Kamang, Sawila, and Wersing.<sup>6</sup> In the Appendix we list only the original reflex, not the secondary development reflected in ‘betel vine’. However, we note that betel vine may be introduced; see also the case of ‘betel nut’ in section 2.1 above.

Initially and medially, \*k remains unchanged in all languages except Adang, where it is reflected as a glottal stop, and Western Pantar, where it is predictably geminated in medial position. Correspondence sets supporting initial \*k include ‘bone’, ‘dog<sub>2</sub>’, ‘fingernail’, and ‘mosquito’. Note that the medial correspondence for Abui *kusɨŋ* ‘fingernail’ is irregular.

Medial \*k is supported by correspondence sets for ‘crouch’, ‘short’, ‘good’, and ‘lizard’. Thus, Nedebug *tuku*, Western Pantar *tuk:a*, Adang *toʔaŋ*, Klón, Kui, Wersing, Teiwa *tuk* < pAP \*tukV ‘short’. The lexeme ‘lizard’ is likely an

<sup>5</sup> It is likely that Malay was only introduced to the Papuan speakers on Alor and Pantar through the Dutch schools that were opened in early 20<sup>th</sup> 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; 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.

<sup>6</sup> The alternation between alveolars and palatals in Adang reflects a phonemic split by which \*d, \*t, and \*n have been palatalized following a vowel sequence ending in a high front vowel, as in ‘betel vine’ (Robinson & Haan 2014). Klón has non-phonemic palatalization in the same environment, while the closely related language Kabola (not in our sample) does not undergo palatalization.

Austronesian borrowing (cf. Alorese *take*), perhaps explaining the anomalous reflexes Adang *tekɔ* and Kamang *tak:ɛ*, the latter of which has an unexpected geminate. However, this form is geminated as expected in Western Pantar *tak:e*.

Final \*k is retained in all languages except Blagar, Adang, Sawila, and Wersing, where it is lost entirely. Only two correspondence sets, ‘one’ and ‘horn’, support this reconstruction. Neither set has 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 Wersing). 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 Western Pantar to be diagnostic in this regard, since Western Pantar geminates do not occur word-finally. This criterion is admittedly problematic, since it is entirely possible that vowel epenthesis preceded gemination in Western Pantar. Furthermore, Western Pantar sometimes lacks cognates for relevant lexical items, as with ‘horn’.

The reconstruction of \*q is supported by the presence of a post-coronal voiceless obstruent phoneme distinct from the velar stop in three Pantar languages. In Teiwa and Nede bang this is a uvular stop; in Kaera a velar fricative. Elsewhere, initial \*q is reflected as /k/, with the exception of Adang, which has glottal stop, and Blagar, which shows both glottal stop and velar stop reflexes. Initial \*q is found in correspondence sets for ‘spear’, ‘itchy’, and ‘tens’. Blagar 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, e.g., Teiwa *qar nuk* ‘ten’.

The medial reflexes of \*q are similar to those in initial position, except that Adang shows loss of medial \*q. Correspondence sets supporting medial \*q include ‘two’, ‘itchy’, ‘white’, and ‘black’. Adang *kak* ‘itchy’ is anomalous, as it retains the medial consonant. Blagar *madzaka* ‘white’ is in fact cognate due to a regular process of glide insertion between the vowels /i/ and /a/, followed by glide fortition: \*miaqa > *miaka* > *mijaka* > *madzaka*. The most interesting reflex of medial \*q is found in Western Pantar. Unlike the other voiceless stops, the uvular stop is not reflected as a geminate in Western Pantar but instead as a non-geminate velar stop. In this regard Western Pantar patterns with the other Pantar languages in distinguishing reflexes of \*q and \*k.

In particular, \*q provides an additional source for non-geminate intervocalic voiceless velar stops in Western Pantar. This in turn may inform reconstruction of final vowels in pAP. Since \*q does not geminate in Western Pantar, Western Pantar *alaku* ‘two’ can readily be derived from \*raqu, supporting reconstruction

of the final vowel. On the other hand, Western Pantar *anuku* ‘one’ corresponds to Tw and Nd forms with velar stops, hence the reconstruction of pAP ‘one’ must contain a velar, not a uvular. The fact that Western Pantar *anuku* does not contain a geminate means that either it has been borrowed or that the vowel has been added following the gemination process. In the absence of any evidence for borrowing we reconstruct \**nuk* ‘one’ without a final vowel.

The evidence for \**q* in final position is extremely limited. One possible example is ‘smoke’, whose correspondences are similar to those for medial \**q* (Kaera *banax* and Wersing *punak*). However, the Teiwa, Nedebang, and Western Pantar reflexes are zero. Another candidate correspondence is ‘rice’: Western Pantar *ala* and Klon, Kui, Wersing *arak*, which compares to Teiwa *qar*, Nedebang *qara*, and Kaera (*na*)*xar*. If the Teiwa, Nedebang, and Kaera forms are interpreted as a result of metathesis of \**r* and \**q*, then this correspondence could also support \**q* in final position, namely, \**araq*. However, we find insufficient evidence to support reconstruction of \**q* in final position.

### 2.3 Fricatives

We reconstruct two fricatives to 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 Adang, and Klon, and strengthens to *t* in Abui, Sawila, and Wersing. In the remaining languages, which include all four Pantar languages in addition to Kui and Kamang, \**s* is retained as *s*. Only in Blagar does the reflex of \**s* exhibit significant variation by position. In initial and final position Blagar has *h* < \**s*, as in Adang and Klon, while in medial position Blagar retains *s* < \**s*. Thus, pAP \**siba* > Blagar *hiba* ‘new’; \**jasi* > Blagar *d3asi* ‘bad’; \**bis* > Blagar *bihi* ‘mat’, with an epenthetic final vowel which was added after the weakening of \**s*. In medial position Nedebang sometimes has as affricate. Thus, \**jiwesin* > Nedebang *jisin* ‘five’, but \**jasi* > Nedebang *jetfi* ‘bad’.

Initial correspondence sets for \**s* are found in ‘new’, ‘wind’, and ‘shark’. Thus, Western Pantar *sab:a*, Blagar *hiba*, Adang *habar*, Klon *həba*, Kui *saba*, Abui *tifa*, Kamang *supa(ka)*, Sawila *tipea*, Wersing *təpa* < \**siba* ‘new’. Medial correspondence sets are found in ‘bad’, ‘fingernail’, ‘tooth’, and ‘five’. Final correspondence sets are found in ‘sit’, ‘stand’, and ‘mat’.

The reconstruction of \**h* is supported by its presence in Western Pantar and Teiwa. The remaining languages lose original \**h*, though Blagar, Adang, and Klon have *h* < \**s*, and Abui has *h* < \**g*. Proto-Alor-Pantar \**h* did not occur in non-initial position. While \**h* is consistently retained in Western Pantar, Teiwa actually has

two reflexes of \*h, the glottal fricative *h* and the pharyngeal fricative *ħ*. This is due to a phonemic split in Teiwa, resulting from an original conditioned distribution where *ħ* occurred only before back vowels, and *h* occurred elsewhere. Modern Teiwa still tends this way, with the pharyngeal fricative generally occurring before back vowels and the glottal fricative preceding front vowels. Klammer (2010) lists only one example of a pharyngeal fricative preceding a front vowel, namely *her* ‘yell, shout, chant, cry aloud’. This form is cognate with Western Pantar *horan*, suggesting that the original form may have contained a back vowel, thus conditioning the Teiwa pharyngeal. This distinction breaks down, however, before low vowels, where a clear synchronic phonemic distinction has developed in Teiwa, as in *haʃan* ‘village’ (< \*haban) vs. *ħaʃ* ‘fish’ (< \*habi).

## 2.4 Nasals

There is a regular and unchanging correspondence of initial /m/ and medial /m/ across all the AP languages. As with the stops, \*m is reflected as a geminate in medial position in Western Pantar. Correspondence sets reflecting initial and medial \*m include ‘come’, ‘betel vine’, ‘sit’, ‘(be) in/on’, ‘fat’, ‘bedbug’, ‘horn’, ‘thatch’, ‘thick’, ‘walk<sub>1</sub>’, and ‘breast’. The reconstruction of pAP initial \*m is thus secure and supported by multiple cognate sets. In medial position subsequent developments may result in nasal-final forms which obey language-specific constraints. For example, Western Pantar does not admit final nasals other than velars, hence Western Pantar *han* ‘breast’ results from later apocope, namely, pAP \*hami > ham:i > ham: > ham > han.

Final \*m is retained as *m* in six of the twelve languages, but only in Teiwa and Kaera does it occur in final position. Blagar and Adang have a velar nasal reflex, while Klön and Kui have an alveolar nasal reflex (Kui *talama* ‘six’ is likely a borrowing from Abui). For many of the languages, forms reflecting final \*m have an epenthetic final vowel, so that the reflex of original final \*m is no longer in final position. Thus, Abui *tala:ma* < \*talam ‘six’, *tama* < \*tam ‘saltwater’. Correspondence sets reflecting final \*m include ‘father’, ‘nose’, ‘six’, and ‘saltwater’. Evidence from ‘saltwater’ weakly supports positing the loss of final \*m in Nedebang.

The behavior of the alveolar nasal mirrors that of the labial nasal in initial and medial position. Proto-Alor-Pantar \*n is retained in all languages and is geminated in medial position in Western Pantar. Correspondence sets supporting initial and medial \*n include ‘one’, 1SG, ‘eat/drink’, ‘smoke’, ‘black’ ‘hold’, ‘give’, ‘die’, ‘ripe’, and ‘name’. Thus, pAP \*nai ‘eat/drink’ > Teiwa, Kaera, Western Pantar, Blagar, Adang *na*, Nedebang *ina*, Klön *na:?*, Kui, Wersing *nai*, Abui, Sawila *ne*; Kamang *ne*.



Final \*n is reflected as a velar nasal in all languages except Teiwa, Klön, and Kui, where it is retained as *n*. The correspondence sets ‘five’, ‘hand/arm’, ‘thatch’, and ‘fingernail’ show irregularities in reflexes of final \*n, perhaps due to borrowing. Nedebang *jisin* ‘five’ has a final alveolar rather than the expected velar and is likely borrowed from neighboring Teiwa *jusan*, while Kaera has *isim* ‘five’ with a final labial nasal, possibly due to influence from the following *tiam* ‘six’ when counting. The correspondence set for ‘fingernail’ is more problematic. Abui *kusiŋ* has the velar nasal as expected but shows an irregular reflex of medial \*s. Western Pantar *kusi* and Klön *kuh* show irregular loss of the final nasal.

## 2.5 Liquids

We reconstruct two liquids \*l and \*r in pAP, though \*r and \*j may have been allophones of a single phoneme in pAP (see Section 2 above). For expository purposes we treat \*r as if it were a phoneme in the present section. There is a relatively regular and unchanging correspondence of initial and medial *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. Correspondences supporting initial \*l include ‘rattan’, ‘crouch’, ‘bark’, ‘walk’, and ‘far’.

Medial \*l is supported by correspondence sets ‘axe’, ‘bathe’, ‘tongue’, and ‘sky’. A few languages show evidence of sporadic \*l > i, for example, Wersing *jebur* ‘tongue’ < \*-lebur. The Pantar languages Teiwa, Nedebang, and Kaera, show irregular loss of medial \*l in ‘six’. Kamang regularly loses \*l between non-front vowels (see Chapter 3), and thus \*talam > Kamang *ta:ma* is expected.

In final position, however, Teiwa, Kaera, and Kamang reflect \*l > i. 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. Adang reflects both *l* and *i* in final position. Synchronically, Adang is losing final *l* among younger speakers and certain dialects, though this only occurs following a sequence of two vowels in final position \*Vil > Vi/\_\_\_# (Robinson & Haan 2014). Further, Nedebang and Western Pantar lose final \*l altogether. Thus, Teiwa *muhui*, Kaera *mogoi*, Western Pantar *mag:i*, Adang *mɔʔɔi*, Klön *məgol*, Kamang *mo:i*, Wersing *mulul* < pAP \*mogol ‘banana’. Other correspondence sets supporting final \*l include ‘child’, ‘bird’, and ‘bat’.

We find insufficient evidence to reconstruct \*r in initial position. In non-initial position pAP clearly distinguished two liquids, and this distinction is preserved in most of the languages. In medial position Nedebang, Western Pantar, and Adang collapse \*l and \*r as *l* (the reflexes in Kamang are less consistent). Abui

reflects \*r > j, represented synchronically as a vowel in final position. The other languages preserve \*r as such. This leaves no direct historical source for r in Nedebang, Western Pantar, and Adang, and we assume that r in these languages has been innovated or diffused from neighboring languages. In modern Western Pantar forms with r are infrequent and do not correspond regularly to other languages. In most cases they reflect lexical innovation, as in Western Pantar *re* ‘bird’ (compare pAP \*(a)dVI). Correspondence sets supporting medial \*r include ‘two’, ‘water’, ‘sing’, ‘bone’, ‘ear’, ‘tail’, and ‘laugh’.

The correspondences for final \*r are similar to those in medial position, except \*r > j (represented here synchronically as a vowel) in Kamang and \*r > Ø in Nedebang and Western Pantar.<sup>7</sup> Adang reflexes of final \*r reflect both l and i, as do its reflexes for final \*l. Correspondence sets supporting final \*r include ‘stone’, ‘scorpion’, ‘lime’, ‘maize’, ‘tongue’, and ‘moon’.

## 2.6 Glides

We reconstruct the two glides \*w and \*j to pAP. In most languages \*w is preserved as w in all positions. In initial position only Blagar *v* and Adang *f* < \*w reflect a change; other languages preserve \*w. Correspondence sets supporting initial \*w include ‘sun’, ‘blood’, ‘stone’, and ‘bathe’. The form ‘blood’ is illustrative, as it has a reflex in every language: Teiwa *wai*, Nedebang *we*, Kaera *we*, Western Pantar *wai*, Blagar *ve*, Adang *foi*, Klön *we?*, Kui *we*, Abui *wea*, Kamang *we*; Sawila *wi*; Wersing *wei*.

We find insufficient evidence to reconstruct \*w in non-initial position. Potential correspondences representing non-initial \*w are likely either to be underlying vowels or to reflect original initial \*w. For example, the root-initial consonant in the word for ‘ear’ is usually analyzed as a glide: Klön *-wer*, Kui *wel*, Abui *wei*, Kamang *wai*, Sawila *-wari*, and Wersing *weri*. However, regardless of the synchronic analyses these forms are likely to reflect an original vocalic form and we reconstruct pAP \*uari. Apparent medial \*w is also found in the word for ‘lime’, Kaera *awar*, Western Pantar *hauwe*, Blagar *avar*, Adang *?afai*, Abui *awai*, Kamang *awoi*. This correspondence matches that for initial \*w and even supports reconstruction 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’, reconstructed as \*jiwesin.

---

<sup>7</sup> Some dialects of Western Pantar have \*r > l in both medial and final position, e.g., Lamma dialect *bat:al* ‘maize’. However, in no dialect of Western Pantar is \*r preserved as r, so forms such as *par* ‘scorpion’ must be borrowings.

Similarly, apparent reflexes of final \*w are better analyzed as reflecting original vowels. For example, Western Pantar *lau*, Adang *loi/lohu*, Abui *lou*, Sawila *lu*, and Wersing *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 Kaera, Blagar, Adang, Kui, Kamang, Sawila, and Wersing the reflex of \*j is lost before a high front vowel [i]. In Western Pantar, it becomes *h* in the same environment. Thus, Teiwa *jas*, Kaera *jas-*, Western Pantar *jasa* < pAP \*jasi ‘bad, broken’; but before a high front vowel Teiwa *jir*, Kaera *ir*, Kamang *ili* < \*jira ‘water’. Correspondence sets supporting initial \*j include ‘water’, ‘bad’, ‘dog<sub>1</sub>’, ‘five’, ‘star’, and ‘laugh’.

In Kui *er* ‘water’ subsequent vowel quality changes have obliterated the environment which triggered loss of \*j. Nedebang and Adang lose the initial syllable of ‘dog<sub>1</sub>’ because the form had final stress and in those languages the initial unstressed syllable was lost. Wersing *wetij* ‘five’ irregularly begins with *w* instead of *j*.

We do not reconstruct \*j in non-initial position. Where non-initial *j* is found in modern languages we assume this is a reflex of a vowel. Examples include Nedebang *buja* ‘betel nut’ < \*bui.

## 2.7 Reconstructed proto-Alor-Pantar vocabulary

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.

Having reconstructed the consonant system we can proceed with a reconstruction of pAP vocabulary. Although we identify 129 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

---

<sup>8</sup> A capital V stands for a vowel, where it is unclear which vowel should be reconstructed.

<sup>9</sup> Several AP languages show medial /g/ or reflexes of medial \*g in ‘laugh’, leading Schapper, Huber & van Engelenhoven (this volume) to reconstruct pAP \*jagir. We find that the correspondences for this medial consonant are highly irregular, and therefore appear to indicate borrowing rather than inheritance. On the other hand, a number of languages unproblematically reflect \*jari, so we reconstruct pAP \*jari as opposed to \*jagir. See the Appendix for a full list of words.

<sup>10</sup> Schapper, Huber & van Engelenhoven (this volume) reconstruct ‘new’ as \*siba(r) with an optional final \*r. In the Timor languages, the final /r/ is found in Makalero. In the modern AP languages, only Adang has a final /r/, but the Adang reflex of pAP \*r is either /l/ or /i/, so we find insufficient evidence to reconstruct ‘new’ with a final \*r at this time.

be found in at least one language of Pantar (Teiwa, Nedebang, Kaera, Western Pantar), one language of West Alor and the Pantar Strait (Blagar, Adang, Klou, Kui), and one language of East Alor (Abui, Kamang, Sawila, Wersing). We exclude from reconstruction very obvious recent borrowings, such as ‘maize’, but we include some forms which are older Austronesian 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 Austronesian (AN) languages in the area.<sup>11</sup>

Table 10 lists 117 vocabulary items which can be reconstructed at the level of pAP on the basis of the correspondence sets above. A full list of the correspondence sets with modern reflexes can be found in the Appendix.

Based on what we know of the phonotactics of the daughter languages, and on the reconstructed pAP vocabulary, we posit a (C)V(C) syllable structure for pAP, with (C)VC or CV(C) as the minimal structure for a single word. In particular, while many of the daughter languages permit words consisting of a single vowel (e.g., Western Pantar *a* ‘tuber’), the reconstructed pAP vocabulary does not contain such forms, although syllables consisting of a single vowel may occur in polysyllabic words. Similarly, while some of the modern languages admit consonant clusters in word-initial onsets, which involve a second liquid second consonant (e.g., Teiwa *bluking* ‘arrow’, Western Pantar *bro* ‘dust’), no consonant clusters are reconstructed for pAP. Underived words in pAP are typically no more than three syllables in length.

### 3 Internal subgrouping

In this section we consider two approaches to explaining the internal relationships of the Alor-Pantar languages. The first is based on the traditional concept of shared phonological innovations. This method robustly identifies shared history, but because the innovations cross-cut one another this method requires subjective weighting of the various innovations. We thus consider also a second less traditional approach based on computational phylogenetics applied to the lexical dataset. Here we apply two methods: split decomposition and Bayesian

---

<sup>11</sup> The fact that these loans can be reconstructed and show regular sound correspondences can be taken as evidence for the claim that the breakup of pAP followed the arrival of AN in the region (perhaps as recently as 3,800-3,000 BP; Spriggs (2011: 511); Pawley (2005: 100)). 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.

statistical techniques. Both methods have been applied successfully to questions of wider family relationships but have only recently been used to explore internal relationships of small language families (e.g. Dunn et al. 2011). In both the traditional and computational approaches we rely on the prior application of the comparative method to establish cognate classes based on regular sound correspondences. That is, we apply these methods to true cognates rather than lexical look-alikes identified based on subjective similarity judgments.

### 3.1 Subgrouping based on shared phonological innovations

The sound correspondences which support reconstruction of the pAP consonant inventory allow us to identify sound changes which have occurred in the daughter languages. While there are many changes which are unique to particular languages, we can identify seventeen sound changes which are each shared by at least two languages (Table 11). Many of these changes are cross-linguistically common, and hence may be of marginal value for subgrouping, for they may have occurred independently in the languages concerned.

Additionally, many of the changes cross-cut each other, further complicating internal subgrouping. For example, the change  $*s > h$  groups Adang with Blagar and Klon, while the change  $*r > l$  groups Adang with Nedebang, Western Pantar, and Abui. This forces a somewhat subjective choice as to which sound change should be given greater weight for the purposes of subgrouping.

The most widespread of these changes is  $*h > \emptyset$ , which occurs in all languages except Teiwa and Western Pantar. 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 Teiwa, Nedebang, and Kaera. This change results in a merger of  $*k$  and  $*q$  in most daughter languages, while Teiwa, Nedebang, and Kaera keep these phonemes distinct. However, closer examination reveals that Western Pantar also distinguishes reflexes of  $*k$  and  $*q$ , though not in all positions. Western Pantar, as noted previously, geminates original stops in medial position, with the exception of  $*q$ . Thus, in medial position the Western Pantar reflexes of  $*k$  and  $*q$  are distinguished as  $k:$  and  $k$ , respectively. Using this evidence to support Western Pantar as maintaining the distinction between  $*k$  and  $*q$  we can then identify a large group of languages which merge these phonemes. The eight languages so identified are precisely the languages of Alor and the Pantar Strait, namely, Blagar, Adang, Klon, Kui, Abui, Kamang, Sawila, and Wersing. We take this change to define a subgroup labeled “Alor.”

Within the Alor group we can distinguish two lower level subgroups. In the east the languages Sawila and Wersing share the innovations \*b > p and \*s > t. The former change is also shared with Kamang; the latter with Abui. So while it is tempting to expand this group, only Sawila and Wersing share both of these innovations, defining a subgroup we refer to as East Alor. In the west the languages Blagar and Adang share innovations \*k > Ø, \*g > ʔ, and \*s > h, defining the Pantar Strait group (labeled “Straits” in the tree). The latter change is also shared with Klou, providing weak support for an intermediate grouping which we label West Alor. The remaining changes cross-cut these and do not provide additional subgrouping information.

The tree based on shared phonological innovations (Figure 1) differs in several ways from previous classifications based on lexicostatistics. In particular, while the eastern languages Sawila and Wersing form a subgroup, they do not constitute primary branches of pAP, as has been suggested in several previous classifications (cf. Wurm 1982; Lewis 2009). This tree has obvious geographic correlates, as shown in Figure 2 below.

The Alor subgroup defined by the merger of pAP velar and uvular stops includes all of the languages of Alor island and the intervening Pantar Straits. The languages of Pantar, with the exception of Blagar which is spoken on both Pantar and in the Straits, do not subgroup together. Within the Alor group are found two primary subgroups: East Alor at the eastern tip of the island, and West Alor comprising the western tip, the Bird’s Head in the Northwest, and the Straits.

### 3.2 Subgrouping based on lexical characters

A second approach to subgrouping delineates subgroups according to shared cognates. For each lexical correspondence set in our data we partitioned the languages into discrete cognate classes. As with the phonological innovations discussed above, the lexical correspondence sets in our data do not all pick out the same subgroups. That is, the cognate sets delineated by some lexical items overlap with those delineated by other lexical items. These overlapping groupings can be visualized in a split graph which represents the distance between the characters in terms of numbers of splits (Figure 3). Details specific to our application of the method are laid out in Robinson & Holton (2012).

Three primary regions can be identified in the graph, each separated by significant reticulation at the center of the graph. An East Alor region groups Kamang, Wersing, and Sawila; a Central Alor region groups Kui, Klou, and Adang; and a Pantar region groups Kaera, Nedebang, Teiwa, and to a lesser extent Western Pantar. The high degree of reticulation within this latter group indicates a strong

2 The internal history of the Alor-Pantar language family

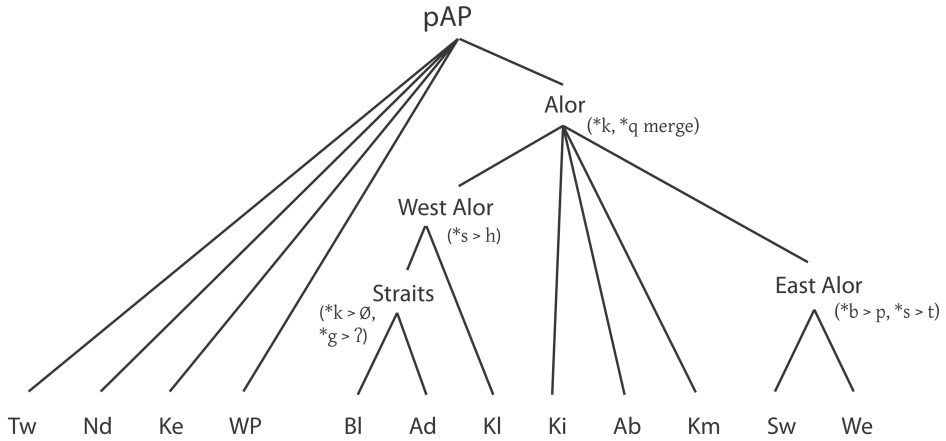


Figure 1: Subgrouping of Alor-Pantar based on shared phonological innovations

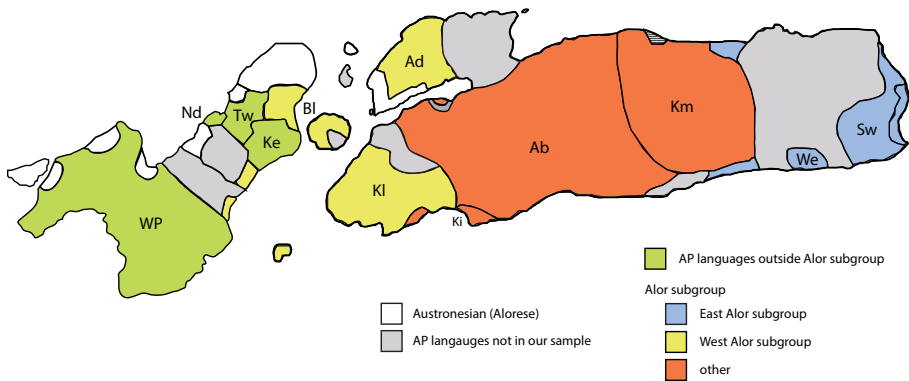


Figure 2: Distribution of subgroups defined based on shared phonological innovations

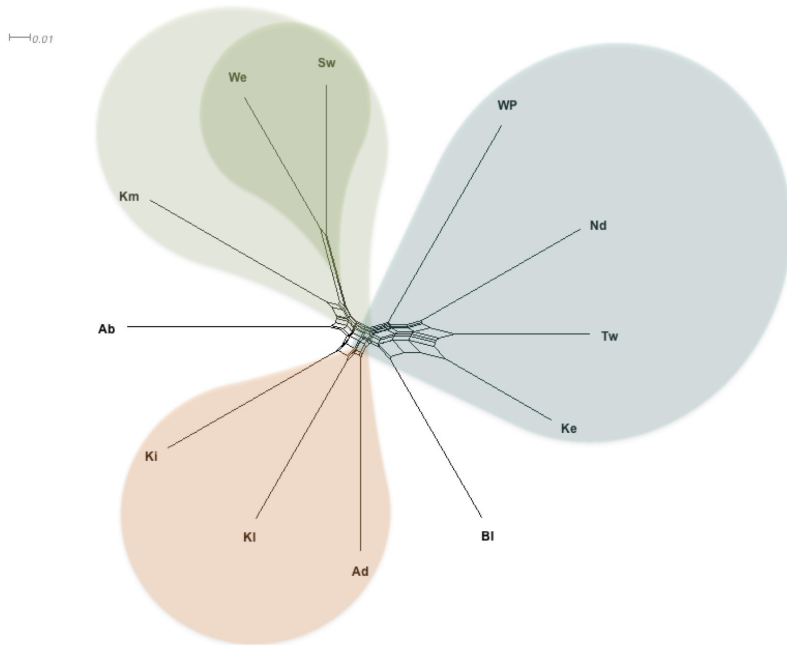


Figure 3: Split graph of lexical character coded into cognate classes, generated using NeighborNet algorithm (Huson & Bryant 2006). pAP node omitted for clarity.

conflicting signal within this region. That is, of these three regions, the Pantar group is particularly non-tree-like, suggesting a pattern of wave-like innovations in this region. In other words, although we found no shared phonological innovations to subgroup these languages together in a traditional tree based on the comparative method (Figure 1), these languages have borrowed a great deal from one another.

A greater degree of reticulation in the graph represents a less tree-like signal in the data. The degree of tree-like signal can be quantified using the delta score metric (Wichmann et al. 2011; Holland et al. 2002). The average delta score for our dataset is a moderately high  $\delta = 0.29$ , reflecting the fact that while some groupings do emerge in Figure 3, there is significant reticulation between those groups. The most tree-like values are found in the East Alor grouping of Kamang, Wersing, and Sawila. The Pantar group of Teiwa, Kaera, and Nedebang has delta scores similar to the mean for the entire dataset; however, the value for West-



## 2 *The internal history of the Alor-Pantar language family*

ern Pantar is significantly higher, suggesting that similarities between Western Pantar and the remainder of the Pantar languages may be due more to borrowing than to shared descent. An unexpected result in the graph in Figure 3 is the position of Blagar as a relative isolate within the family. In contrast to the subgrouping based on the comparative method, Blagar groups not with Adang and Klon but rather with the Pantar languages – and then only weakly so.

A second method of subgrouping based on lexical characters uses Bayesian statistical techniques to search for trees which are most compatible with the cognate classes coded in our data.<sup>12</sup> The results are summarized in Figure 4 as a maximum clade credibility tree. The clade credibility values listed below each node indicate the percentage of sampled trees which are compatible with that node. These values are for the most part either at or near one hundred percent (1.00), indicating that this consensus tree is compatible with almost all of the trees sampled in the analysis. Lower figures appear at exactly those nodes already shown to be problematic via the other subgrouping methods, namely Western Pantar, Abui, and Kamang.

To a large extent the groupings in the Bayesian tree are compatible with those in the split graph. First, Sawila (Sw) and Wersing (We) are shown to be closely related, a grouping which was also present in the classification based on the comparative method (Figure 1). Second, there is a Pantar grouping of Kaera (Ke), Teiwa (Tw), Nedebang (Nd), and Western Pantar (WP). Third, the position of Blagar (Bl) at the highest node coordinate to the Alor languages is consistent with its position in the split graph, though, as noted above, this differs significantly from its position in the tree based on the traditional application of the comparative method (Figure 1). On the other hand, there are also some incompatibilities between the Bayesian tree and the split graph. For example, in the tree based on lexical characters Adang (Ad) and Klon (Kl) are shown forming a group without Kui (Ki), contra both the splits graph and the tree calculated using the comparative method.

---

<sup>12</sup> We employ a Markov Chain Monte Carlo (MCMC) method to search through the probability space of all possible trees, using a relaxed Dollo model. Details of this implementation can be found in Robinson & Holton (2012), which compares the results of several different models, using both MrBayes 3.2.1 (Ronquist & Huelsenbeck 2003) and BEAST 1.7.2 (Drummond et al. 2012), running each model for at least 10 million iterations with a sample rate of 1000 and a burn-in of 25 percent. Each model converged after approximately 1.5 million iterations, and the best performing model (i.e., that with the highest likelihood) was found to be the relaxed Dollo model implemented in BEAST. This model has been argued to be particularly appropriate to linguistic data, since it assumes that innovations may arise only once but may be lost multiple times independently (Pagel 2009).

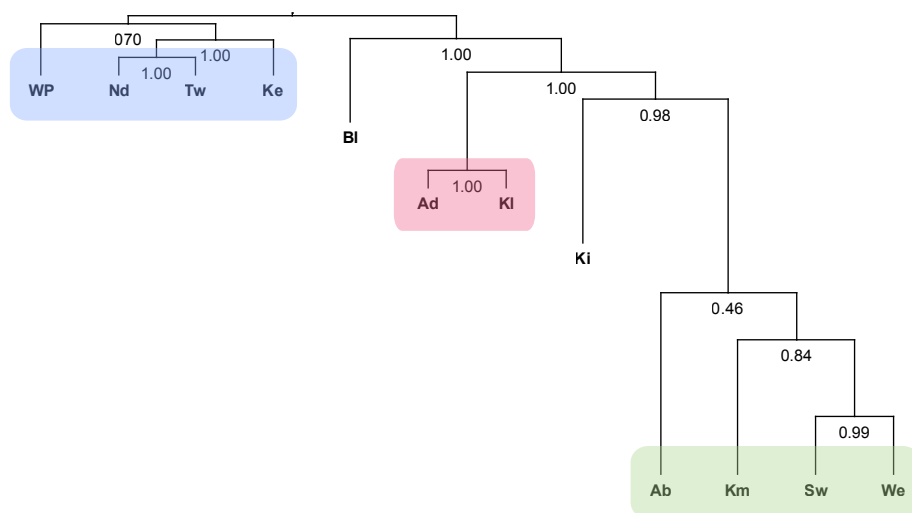


Figure 4: Bayesian MCMC maximum clade credibility tree for lexical data (relaxed Dollo model), with clade credibility values indicated. pAP node omitted

Though not immediately apparent based on visual inspection of the maximum clade credibility tree in Figure 4, the subgrouping based on lexical characters is also largely compatible with that based on phonological innovations. To demonstrate this we repeated the Bayesian analysis with the constraint that all sampled trees be compatible with the subgroups identified by the comparative method, keeping all other parameters constant.<sup>13</sup> We then applied a marginal likelihood analysis to the results of each model, which yielded a Bayes factor of 1.1726, only slightly favoring the constrained model over the unconstrained one.<sup>14</sup> The model based on lexical characters independently identifies the same subgroups found using a completely different methodology based on phonological characters, providing additional support for the robustness of the model. This lends support for those subgroups identified in the model based on lexical characters which are not found in the subgrouping based on phonological innovations. In particular, we have some evidence for the existence of an East Alor subgroup comprised of Abui (Ab), Kamang (Km), Sawila (Sw) and Wersing (We), even though this subgroup is not identified in the tree based on phonological characters.

<sup>13</sup> The authors thank Michael Dunn for suggesting this innovative approach.

<sup>14</sup> Marginal likelihood was estimated using Tracer 1.5 (Rambaut & Drummond 2007).

## 4 Discussion

The examination of sound correspondences across the Papuan languages of Alor and Pantar robustly supports the identification and reconstruction of an Alor-Pantar family. Our comparative work also allows us to propose internal subgroups within Alor-Pantar, but the overall linguistic picture is extremely complex, defying a model based solely on inheritance. Widespread multilingualism is the norm in the region, and borrowings from neighboring languages – such as Western Pantar *bagis* ‘whine’ from Deing *bagis* ‘cry’ – are extremely common.<sup>15</sup> Additionally, genetic studies indicate that East Nusantara, and the Alor-Pantar region in particular, is a melting pot with a long history of admixture (Mona et al. 2009), and it may well be that an analogous situation holds for languages, reflecting extensive borrowing and metatypy. Thus it is not surprising that different methods reveal different trees for the family.

The family tree based on phonological innovations identified by the comparative method (Figure 1) shows the highest level of diversity on Pantar, suggesting the languages originated in Pantar, spreading east. The tree based on the lexical characters (Figure 3) suggests that the languages of Alor originated in the Pantar Strait (around the area where Blagar is spoken today) with subsequent migration eastward. These two trees reveal different aspects of the prehistory of the AP languages. Phonological innovations show the greatest degree of diversity on Pantar, suggesting a long history of settlement there. Lexical innovations are compatible with an original settlement in the Pantar Strait. We propose that the original settlement was indeed in the area of the Pantar Strait with a very early split towards Pantar. That early settlement of Pantar led to the diversity we see there in terms of phonological innovations. The lexical innovations show less diversity on Pantar due to subsequent diffusion (as indicated by the significant reticulation for the languages of Pantar in Figure 3). As languages spread eastward from the Pantar Strait into Alor, new lexical innovations were restricted to smaller and smaller subgroups in the east, leading to the embedded structure in the tree based on lexical characters (Figure 4). However, the Pantar Strait languages (particularly Blagar and Adang) constitute a more recent linguistic area across which phonological innovations have been shared, leading to their close subgrouping in the tree based on phonological innovations (Figure 1).

While this picture is fairly complex in terms of layers of history, it is not unexpected in a region of significant warfare and shifting alliances overlaid by sev-

---

<sup>15</sup> Geographically, Deing lies between Western Pantar and Teiwa. It appears to be closely related to Teiwa.

eral periods of contact from different outside groups (first the ancestors of today's Muslim speakers of Alorese, then the Dutch, and now Indonesian). A more complete picture of the prehistory of the region must await evidence from other disciplines, particularly archaeology and genetics.

## **Appendix**

### **Cognate sets**

Here we list 129 cognate sets reflecting regular sound correspondences. There are only 127 distinct meanings, as two of the meanings, 'dog' and 'walk', are found in more than one cognate set; these are indicated with subscripts following the gloss. In the table the correspondence sets are listed alphabetically by English gloss. Languages are arranged in order roughly from west to east with the western-most languages on the left and the eastern-most languages on the right. Correspondence sets may include irregular forms when they serve to demonstrate the correspondence under discussion. In these cases the irregular forms are denoted with a preceding double dagger (‡). We reconstruct pAP forms only when we have broad geographic support in minimally one language of Pantar (Teiwa, Nedebang, Kaera, Western Pantar), one language of West Alor and the Pantar Strait (Blagar, Adang, Klou, Kui), and one language of East Alor (Abui, Kamang, Sawila, Wersing). Of these 129 correspondences, 117 reconstruct to the level of pAP.

2 *The internal history of the Alor-Pantar language family*

Table 7: Alor-Pantar nasal correspondences

| pAP | Env | Tw | Nd | Ke | WP | Bl | Ad | Kl | Ki | Ab | Km | Sw | We |
|-----|-----|----|----|----|----|----|----|----|----|----|----|----|----|
| *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  | ∅  | 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 8: Alor-Pantar liquid correspondences

| pAP | Env | Tw | Nd | Ke | WP | Bl | Ad  | Kl | Ki | Ab | Km  | Sw | We |
|-----|-----|----|----|----|----|----|-----|----|----|----|-----|----|----|
| *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  | ∅  | i  | ∅  | l  | l/i | l  | l  | l  | i   | l  | l  |
| *r  | V_V | r  | l  | r  | l  | r  | l   | r  | r  | j  | l   | r  | r  |
| *r  | _#  | r  | ∅  | r  | ∅  | r  | l/i | r  | r  | i  | i   | r  | r  |

Table 9: Alor-Pantar glide correspondences

| pAP | Env | Tw | Nd | Ke | WP | Bl | Ad | Kl | Ki | Ab | Km | Sw | We |
|-----|-----|----|----|----|----|----|----|----|----|----|----|----|----|
| *w  | #_  | w  | w  | w  | w  | v  | f  | w  | w  | w  | w  | w  | w  |
| *j  | #_  | j  | j  | j  | j  | dʒ | s  | ∅  | j  | j  | j  | j  | j  |

Table 10: Reconstructed pAP vocabulary

|                                     |                                    |                                   |
|-------------------------------------|------------------------------------|-----------------------------------|
| * <i>(a)dVl</i> <sup>8</sup> ‘bird’ | * <i>jari</i> <sup>9</sup> ‘laugh’ | * <i>por</i> ‘dry in sun’         |
| * <i>en(i,u)</i> ‘name’             | * <i>jasi</i> ‘bad, broken’        | * <i>p{i,u}nV</i> ‘hold’          |
| * <i>aman</i> ‘thatch’              | * <i>jibV</i> ‘star’               | * <i>pi-</i> ‘1PL.INCL’           |
| * <i>aqana</i> ‘black’              | * <i>jibar</i> ‘dog’               | * <i>purVn</i> ‘spit’             |
| * <i>-ar</i> ‘vagina’               | * <i>jira</i> ‘water’              | * <i>pVr</i> ‘scorpion’           |
| * <i>araq</i> ‘two’                 | * <i>jira(n)</i> ‘fly’ (v.)        | * <i>rVsi</i> ‘goanna’            |
| * <i>-asi</i> ‘bite’                | * <i>jiwesin</i> ‘five’            | * <i>qaba(k)</i> ‘spear’          |
| * <i>bagai</i> ‘crocodile’          | * <i>kin</i> ‘mosquito’            | * <i>qar-</i> ‘tens’              |
| * <i>bagori</i> ‘yellow’            | * <i>kusin</i> ‘fingernail’        | * <i>siba</i> <sup>10</sup> ‘new’ |
| * <i>baj</i> ‘pig’                  | * <i>kVt</i> ‘flea’                | * <i>sib(a,i)r</i> ‘shark’        |
| * <i>-bat</i> ‘leg’                 | * <i>lam(ar)</i> ‘walk’            | * <i>talam</i> ‘six’              |
| * <i>bis</i> ‘mat’                  | * <i>-lebur</i> ‘tongue’           | * <i>tam</i> ‘saltwater’          |
| * <i>bob</i> ‘wave’                 | * <i>lete</i> ‘far’                | * <i>tama</i> ‘fat’               |
| * <i>bui</i> ‘betel nut’            | * <i>luk(V)</i> ‘crouch’           | * <i>-tan</i> ‘hand/arm’          |
| * <i>bukan</i> ‘guard’              | * <i>IVu</i> ‘bark’(v.)            | * <i>tapai</i> ‘pierce’           |
| * <i>bunaq</i> ‘smoke’              | * <i>madel</i> ‘bat’               | * <i>tas</i> ‘stand’              |
| * <i>dar(a)</i> ‘sing’              | * <i>magi</i> ‘hear’               | * <i>tei</i> ‘tree’               |
| * <i>dul(a)</i> ‘slippery’          | * <i>mai</i> ‘come’                | * <i>temek</i> ‘bedbug’           |
| * <i>dumV</i> ‘thick’               | * <i>mait</i> ‘betel vine’         | * <i>tena</i> ‘ripe’              |
| * <i>dur</i> ‘rat’                  | * <i>-mam</i> ‘father’             | * <i>-ten</i> ‘wake s.o.’         |
| * <i>ede</i> ‘burn’                 | * <i>mari</i> ‘bamboo’             | * <i>tia</i> ‘recline’            |
| * <i>-ena</i> ‘give’                | * <i>mi</i> ‘(be) in/on’           | * <i>tia</i> ‘expel’              |
| * <i>ga-</i> ‘3SG’                  | * <i>mid</i> ‘climb’               | * <i>-tiari(n)</i> ‘close’ (v.)   |
| * <i>ge-</i> ‘3GEN’                 | * <i>-mim</i> ‘nose’               | * <i>-tok</i> ‘stomach’           |
| * <i>gi-</i> ‘3PL’                  | * <i>min(a)</i> ‘die’              | * <i>tukV</i> ‘short’             |
| * <i>ha-</i> ‘2SG’                  | * <i>mis</i> ‘sit’                 | * <i>-uaqal</i> ‘child’           |
| * <i>habi</i> ‘fish’                | * <i>mogol</i> ‘banana’            | * <i>-uari</i> ‘ear’              |
| * <i>haban</i> ‘village’            | * <i>mudi</i> ‘body hair’          | * <i>uas</i> ‘tooth’              |
| * <i>hada</i> ‘fire, firewood’      | * <i>mudin</i> ‘plant’ (v.)        | * <i>uku</i> ‘knee’               |
| * <i>hagur</i> ‘yawn’               | * <i>-muk</i> ‘horn’               | * <i>-wa</i> ‘mouth’              |
| * <i>hami</i> ‘breast’              | * <i>mVn</i> ‘rotten’              | * <i>wadi</i> ‘sun’               |
| * <i>has</i> ‘excrement’            | * <i>na-</i> ‘1SG’                 | * <i>wai</i> ‘blood’              |
| * <i>hasak</i> ‘empty’              | * <i>nai</i> ‘eat/drink’           | * <i>wai</i> ‘roof’               |
| * <i>hawar</i> ‘lime’               | * <i>nan(a)</i> ‘sibling (older)’  | * <i>war</i> ‘stone’              |
| * <i>hipar</i> ‘dream’              | * <i>nuk</i> ‘one’                 | * <i>wata</i> ‘coconut’           |
| * <i>hu:ba</i> ‘sugarcane’          | * <i>oda</i> ‘throw’               | * <i>weli</i> ‘bathe’             |
| * <i>is(i)</i> ‘fruit’              | * <i>-ora</i> ‘tail’               | * <i>wur</i> ‘moon’               |

2 *The internal history of the Alor-Pantar language family*

Table 11: Sound changes found in at least two languages

| Change   | Languages  |
|----------|--|
| *b>f     | Teiwa, Nedebang, Abui (in Teiwa and Nedebang only non-initially)                     |
| *b>p     | Kamang, Sawila, Wersing  |
| *d>r     | Abui, Kui (in Kui only finally)  |
| *g>ʔ     | Blagar, Adang  |
| *k>Ø/_#  | Blagar, Adang  |
| *q>k     | Western Pantar, Blagar, Adang (ʔ < k < *q), Klön, Kui, Abui, Kamang, Sawila, Wersing |
| *s>h     | Blagar, Adang, Klön  |
| *s>t     | Abui, Sawila, Wersing  |
| *h>Ø     | everywhere but Teiwa and Western Pantar  |
| *m>ŋ/_#  | Western Pantar, Blagar, Adang  |
| *n>ŋ/_#  | Nedebang, Kaera, Western Pantar, Blagar, Adang, Abui, Kamang, Sawila, Wersing        |
| *l>i/_#  | Teiwa, Kaera, Adang, Kamang  |
| *l>Ø/_#  | Nedebang, Western Pantar, Abui   |
| *r>l/V_V | Nedebang, Western Pantar, Adang, Kamang  |
| *r>Ø/_#  | Teiwa, Kaera, Western Pantar   |
| *r>i/_#  | Blagar, Kui, Abui  |

| Gloss              | pAP                 | Tw            | Nd           | Ke            | WP              | Bl             | Ad              | Kl             | Ki                     | Ab                       | Kim          | Sw                | We              |
|--------------------|---------------------|---------------|--------------|---------------|-----------------|----------------|-----------------|----------------|------------------------|--------------------------|--------------|-------------------|-----------------|
| 'axe' <sup>1</sup> |                     | jas           | jetfi        | jas-          | <i>baliŋ</i>    |                | <i>baliŋ</i>    |                |                        | <i>faliŋ</i>             | <i>paliŋ</i> |                   |                 |
| 'bad, broken'      | *jasi               | jas           | jetfi        | jas-          | <i>jas</i>      | <i>dʒasi</i>   | <i>sah</i>      | <i>ja-h</i>    |                        |                          |              | <i>ja-ti</i>      |                 |
| 'bamboo'           | *mari               |               |              |               | <i>mali</i>     | <i>mari</i>    | <i>mai</i>      | <i>(du)mar</i> |                        | <i>mai</i>               | <i>mai</i>   |                   |                 |
| 'banana'           | *mogol              | <i>muhui</i>  | <i>ʒmaj</i>  | <i>mogoi</i>  | <i>mag-i</i>    | <i>ʒmɔl</i>    | <i>mɔʔoi</i>    | <i>məgɔl</i>   |                        |                          | <i>moi</i>   |                   | <i>mulul</i>    |
| 'bark' (v.)        | *Ivu                |               |              |               | <i>lau</i>      | <i>ʒorow</i>   | <i>lou</i>      |                |                        | <i>loi</i>               |              | <i>lu</i>         | <i>aloi</i>     |
| 'bat'              | *madel              | <i>madi</i>   | <i>ʒmara</i> | <i>ʒmerei</i> | <i>mad-e</i>    | <i>demeɬ</i>   | <i>ʒmadiruw</i> | <i>mədel</i>   | <i>madel</i>           | <i>marel</i>             | <i>matei</i> | <i>ʒmadi:(ku)</i> | <i>ʒmudu(k)</i> |
| 'bathe'            | *weli               | <i>wei</i>    |              | <i>temek</i>  |                 | <i>wela</i>    | <i>foil</i>     | <i>wel</i>     | <i>weli</i>            | <i>-wel</i>              | <i>-wei</i>  | <i>wile</i>       | <i>-weli</i>    |
| 'bedbug'           | *temek              |               |              |               |                 | <i>teme</i>    | <i>ʒtame?</i>   | <i>tamek</i>   |                        | <i>tameki</i>            |              |                   | <i>mekiʔ</i>    |
| 'betel nut'        | *bui                | <i>bui</i>    | <i>buja</i>  | <i>bui</i>    | <i>bu</i>       | <i>bu</i>      | <i>bu</i>       | <i>bui</i>     | <i>bui</i>             | <i>fu</i>                |              | <i>pu</i>         | <i>pui</i>      |
| 'betel vine'       | *mait               | <i>met</i>    | <i>mata</i>  | <i>mat</i>    | <i>meta</i>     | <i>mat</i>     | <i>metf</i>     | <i>meh</i>     | <i>mesin</i>           | <i>me:tiŋ</i>            | <i>maisi</i> | <i>masi</i>       | <i>mas</i>      |
| 'bird'             | *(a)dVl             | <i>dai</i>    | <i>ʒdaja</i> |               | <i>ʒduŋ</i>     |                |                 | <i>adol</i>    | <i>adol</i>            | <i>ruwol<sup>h</sup></i> | <i>atoi</i>  | <i>adala</i>      | <i>adol</i>     |
| 'bite'             | *asi                | <i>si</i>     | <i>tʃia</i>  | <i>si:</i>    | <i>sia</i>      |                | <i>-eh</i>      | <i>eh</i>      | <i>-es</i>             |                          | <i>-eh</i>   |                   |                 |
| 'black'            | *aqana              | <i>qaʔan</i>  | <i>qana</i>  | <i>xan</i>    | <i>ʒana</i>     | <i>kaʔana</i>  | <i>(l)qaʔan</i> | <i>akan</i>    | <i>akana</i>           | <i>akan</i>              |              | <i>akana</i>      | <i>akeŋ</i>     |
| 'blood'            | *wai                | <i>wai</i>    | <i>we</i>    | <i>wei</i>    | <i>wai</i>      | <i>we</i>      | <i>foi</i>      | <i>we?</i>     | <i>we</i>              | <i>wea</i>               | <i>we:</i>   | <i>wi:</i>        | <i>wei</i>      |
| 'body hair'        | *mudi               | <i>mud</i>    | <i>mudi</i>  | <i>-mudu</i>  | <i>-mudi</i>    |                | <i>amudi</i>    | <i>amudi</i>   |                        | <i>amur</i>              |              | <i>madi</i>       | <i>mudi</i>     |
| 'bone'             |                     | <i>kir</i>    | <i>kili</i>  | <i>kiri</i>   | <i>kira</i>     |                |                 |                |                        |                          | <i>ami</i>   | <i>-a:mi</i>      | <i>ami</i>      |
| 'breast'           | *hami               | <i>-ham</i>   | <i>ami</i>   |               | <i>-haŋ</i>     |                |                 |                |                        |                          |              |                   |                 |
| 'burn'             | *ede                | <i>de?</i>    |              | <i>de</i>     | <i>ʔede</i>     |                |                 |                |                        | <i>ʒdiei</i>             |              |                   | <i>ʒol</i>      |
| 'child'            | *uaqal              | <i>-oqai</i>  | <i>uaqa</i>  | <i>-uax</i>   | <i>wake</i>     | <i>-oal</i>    | <i>ʔai</i>      | <i>ʒul</i>     | <i>ʒol</i>             |                          |              | <i>mada</i>       | <i>ʒol</i>      |
| 'climb'            | *mid                | <i>mir</i>    |              |               | <i>midan</i>    |                | <i>mid</i>      | <i>mid</i>     |                        |                          |              |                   |                 |
| 'close' (v.)       | *-tiar(n)           |               |              | <i>teriŋ</i>  | <i>ʒtiariŋ</i>  | <i>teriŋ</i>   | <i>tel</i>      | <i>(u)ter</i>  | <i>(u)teri</i>         | <i>wata</i>              | <i>wate</i>  | <i>-ti:ra</i>     | <i>(le)ter</i>  |
| 'coconut'          | *wata               | <i>wat</i>    | <i>wata</i>  | <i>wat</i>    | <i>wata</i>     | <i>wet</i>     | <i>fa</i>       | <i>ʒata</i>    | <i>ʒbat</i>            | <i>wata</i>              | <i>wate</i>  | <i>wata</i>       | <i>wata</i>     |
| 'come'             | *mai <sup>5</sup>   | <i>ma</i>     | <i>ma</i>    | <i>ma</i>     | <i>ma</i>       | <i>ma</i>      | <i>ma</i>       | <i>ma</i>      | <i>mai</i>             | <i>me</i>                |              | <i>me</i>         | <i>amai</i>     |
| 'crocodile'        | *bagai              | <i>bahai</i>  |              | <i>bagai</i>  | <i>ʒbagai</i>   |                | <i>baʔai</i>    | <i>bəgai</i>   | <i>ʒbuai</i>           | <i>fahai</i>             | <i>pie:</i>  |                   | <i>luku(k)</i>  |
| 'crouch'           | *luk(V)             |               |              |               | <i>luk:itiŋ</i> |                |                 | <i>luk</i>     | <i>luk<sup>6</sup></i> | <i>luk<sup>7</sup></i>   |              |                   |                 |
| 'die'              | *min(a)             | <i>min</i>    | <i>mina</i>  | <i>nimin</i>  | <i>ʒhin-a</i>   | <i>mina</i>    | <i>min</i>      | <i>min</i>     | <i>min</i>             | <i>mon</i>               |              |                   |                 |
| 'dog'              | *jibar <sup>8</sup> | <i>ji'var</i> | <i>ʒbar</i>  | <i>i'bar</i>  | <i>ja'be</i>    | <i>dʒa'bar</i> | <i>bel</i>      |                |                        |                          |              |                   |                 |



2 The internal history of the Alor-Pantar language family

| Gloss                     | pAP                  | Tw            | Nd             | Ke            | WP                         | Bl             | Ad                     | Kl           | Ki            | Ab              | Km             | Sw                        | We                        |
|---------------------------|----------------------|---------------|----------------|---------------|----------------------------|----------------|------------------------|--------------|---------------|-----------------|----------------|---------------------------|---------------------------|
| 'dog <sub>2</sub> '       |                      |               |                |               |                            |                |                        | <i>kur</i>   | <i>kur</i>    | <i>ka:i</i>     | <i>kui</i>     |                           |                           |
| 'dream'                   | *hipar               |               |                | <i>ipar</i>   | <i>hipe</i>                | <i>ipar</i>    | <i>apai</i>            | <i>eper</i>  |               | <i>piei</i>     | <i>-foi</i>    |                           |                           |
| 'dry in sun'              | *por                 |               |                | <i>poriŋ</i>  | <i>ɕpuariŋ</i>             | <i>poriŋ</i>   | <i>poiŋ</i>            | <i>upu:r</i> |               |                 |                | <i>po:por<sup>9</sup></i> |                           |
| 'dry'                     |                      |               |                |               |                            | <i>ɕtaʔata</i> | <i>taʔat</i>           | <i>təkət</i> | <i>takata</i> | <i>takat</i>    |                |                           |                           |
| 'ear'                     | *uari                | <i>uar</i>    | <i>ow</i>      | <i>uar</i>    | <i>uwe</i>                 | <i>veri</i>    | <i>-fel</i>            | <i>wer</i>   | <i>ɕwel</i>   | <i>wei</i>      | <i>wai</i>     | <i>-wari</i>              | <i>weri</i>               |
| 'eat/drink' <sup>10</sup> | *nai                 | <i>na</i>     | <i>ina</i>     | <i>na</i>     | <i>na</i>                  | <i>na</i>      | <i>na</i>              | <i>na:ʔ</i>  | <i>nai</i>    | <i>ne:</i>      | <i>ne:</i>     | <i>ne:</i>                | <i>nai</i>                |
| 'empty'                   | *hasak               | <i>hasak</i>  |                | <i>isik</i>   | <i>hak:as<sup>11</sup></i> |                |                        |              | <i>taka</i>   | <i>taka</i>     | <i>saka</i>    |                           |                           |
| 'excrement'               | *has                 | <i>has</i>    |                | <i>ter</i>    | <i>has</i>                 | <i>ɕ:a:s</i>   | <i>ah</i>              | <i>ihi</i>   | <i>es</i>     | <i>ɕ:asi</i>    | <i>asi</i>     | <i>atu</i>                | <i>atu</i>                |
| 'expel'                   | *tiara               | <i>-tiar</i>  | <i>-tiala</i>  | <i>ter</i>    |                            | <i>-teri</i>   | <i>(ate)kel</i>        | <i>let</i>   |               |                 | <i>ti:ra</i>   | <i>-(pan)ter</i>          |                           |
| 'far'                     | *lete <sup>12</sup>  |               |                |               |                            |                | <i>let</i>             | <i>let</i>   |               |                 | <i>letei</i>   |                           |                           |
| 'fat'                     | *tama                | <i>tama?</i>  |                | <i>tama</i>   |                            | <i>tama</i>    | <i>tama(r) tɔma(d)</i> | <i>let</i>   | <i>tama</i>   | <i>tama(da)</i> |                |                           |                           |
| 'father'                  | *mam                 |               |                | <i>-mam</i>   |                            | <i>-maŋ</i>    | <i>-maŋ</i>            | <i>-man</i>  | <i>ɕ-ma</i>   | <i>mama</i>     |                |                           |                           |
| 'fingernail'              | *kusin               |               | <i>kufiŋ</i>   | <i>kusiŋ</i>  | <i>ɕ:kusi</i>              | <i>ɕ:kusiŋ</i> | <i>ɕ:huin</i>          | <i>ɕ:kuh</i> | <i>kusin</i>  | <i>ɕ:kusiŋ</i>  | <i>kusiŋ</i>   |                           |                           |
| 'fire'                    | *hada                | <i>har</i>    | <i>ar</i>      | <i>ad</i>     | <i>had:i<sup>13</sup></i>  | <i>ʔad</i>     | <i>ɕ:a</i>             | <i>adaʔ</i>  | <i>ar</i>     | <i>ara</i>      | <i>ati</i>     | <i>ada</i>                | <i>ada</i>                |
| 'fish'                    | *habi                | <i>haɕ</i>    | <i>a:fi</i>    | <i>ab</i>     | <i>hap</i>                 | <i>a:b</i>     | <i>ab</i>              | <i>ibiʔ</i>  | <i>eb</i>     | <i>afu</i>      | <i>api</i>     | <i>api</i>                | <i>api</i>                |
| 'five'                    | *jiwesin             | <i>jusan</i>  | <i>ɕ:jisin</i> | <i>isim</i>   | <i>jasitj</i>              | <i>ɕ:isitj</i> | <i>ifihiŋ</i>          | <i>εweh</i>  | <i>jesan</i>  | <i>jetitj</i>   | <i>iwesitj</i> | <i>jo:itj</i>             | <i>wetitj</i>             |
| 'flea'                    | *kVt                 | <i>ɕ:hat</i>  |                | <i>kati</i>   |                            | <i>ɕ:ʔut</i>   |                        | <i>kot</i>   |               |                 |                |                           | <i>tokoʔ<sup>14</sup></i> |
| 'fly' (v.)                | *jira(n)             | <i>jir-an</i> | <i>jila</i>    | <i>ir</i>     | <i>hil:an</i>              |                |                        |              |               |                 |                | <i>iritj</i>              | <i>ire</i>                |
| 'fruit'                   | *is(i) <sup>15</sup> | <i>jis</i>    | <i>itfi</i>    | <i>isi</i>    | <i>ɕ:hisa</i>              | <i>(?)hi</i>   | <i>ih</i>              | <i>ih</i>    |               |                 | <i>ih</i>      | <i>ɕ:si</i>               | <i>ɕ:is<sup>16</sup></i>  |
| 'garden'                  |                      | <i>mahar</i>  | <i>maxara</i>  | <i>mag:ar</i> |                            | <i>maʔad</i>   |                        |              |               |                 |                |                           |                           |
| 'give'                    | *-enV                | <i>-an</i>    | <i>-ena</i>    | <i>-eŋ</i>    | <i>-nu</i>                 | <i>-enaj</i>   | <i>-en</i>             | <i>-en</i>   | <i>-ana</i>   |                 | <i>-n</i>      |                           | <i>-eni(r)</i>            |
| 'good'                    |                      |               |                |               |                            | <i>noʔ</i>     | <i>nok</i>             | <i>nok</i>   |               |                 |                |                           |                           |
| 'guard'                   |                      | <i>ɕbohon</i> |                | <i>bukan</i>  | <i>baukan</i>              |                | <i>buk</i>             | <i>buk</i>   |               |                 | <i>-pulkan</i> |                           |                           |
| 'hand/arm'                | *tan                 | <i>-tan</i>   | <i>-tan</i>    | <i>tan</i>    | <i>t:an</i>                | <i>tan</i>     | <i>tan</i>             | <i>tan</i>   | <i>tan</i>    | <i>tan</i>      | <i>tan</i>     | <i>tan</i>                | <i>teŋ</i>                |

| Gloss                  | pAP                 | Tw               | Nd        | Ke      | WP                  | Bl       | Ad       | Kl      | Ki      | Ab    | Km       | Sw                   | We                 |
|------------------------|---------------------|------------------|-----------|---------|---------------------|----------|----------|---------|---------|-------|----------|----------------------|--------------------|
| 'hear'                 | *magi <sup>17</sup> | tuta(h)          | tutu-     | tutu(k) | tutu                | me?e     | ma?eh    | magih   | magi    | mahi  | mai      | maji:ŋ               |                    |
| 'hearth'               |                     | pin              | pini      | pin     | pini                | pina     | puin     | puin    | puna    | pun   | fun      | puni <sup>19</sup>   | poiŋ <sup>20</sup> |
| 'hold' <sup>18</sup>   | *p{t,u}nV           |                  |           | muk     |                     | mu       | mu       | muk     | muk     | muk   | ɸmu:     |                      |                    |
| 'horn'                 | *muk                |                  |           | mi      | me                  | mi       | mi       | mi      | mi      | mi    | mi       | ma                   |                    |
| '(be) in/on'           | *mi                 | me?              |           | mi      | me                  | mi       | mi       | mi      | mi      | mi    | mi       |                      |                    |
| 'itchy'                |                     | qa:q             | qaga      | xaxaw   | kaka                | kaka     | ɸkak     | ka:k    |         |       |          |                      |                    |
| 'knee'                 | *uku                | ku:?             | uku       | uku     | uk:a(t)             | (k)uku   |          | -uk     | -uk     |       |          |                      |                    |
| 'laugh'                | *jari               | ɸjahar           | ɸgela     | ɸogar   | jali                | asal     | ɸ?ogar   | jeri    |         |       | ɸje:i    | (tasur)uku (seseb)uk | jer                |
| 'leg'                  | *-bat <sup>21</sup> | -ɸat             |           | -bat    | -uta                | -(-e?)fa |          |         |         |       |          |                      |                    |
| 'lime'                 | *hawar              | hor              | wa        | awar    | hawwe               | avar     | ?afai    | ewer    | or      | awai  | awoi     |                      | or                 |
| 'lizard'               |                     | takok taka(ra:b) |           | tek     | tak:e               | teke     | ɸteko    | takok   | takok   | tekok | ɸtak:e   | tako                 |                    |
| 'maize'                |                     | batar            | ba:ta     | batar   | bat:e               | batar    | batε     | bat     | batar   | fat   | patei    | patara               | peter              |
| 'mat'                  | *bis                | ɸbi:             |           | bis     | bis                 | bili     | buh      | ur      | ur      | ur    | ur       | ɸbusi                | ɸbiti?             |
| 'moon'                 | *wur                | wur              | hula      | ur      | uru                 | uru      | ul       | ur      | ur      | ur    | wui      | ura(k)               |                    |
| 'mosquito'             | *kin                | ki?m             | kim(balu) | kiŋ     | ki?2                | kini     | ?m       | ikin    | kin     | kin   | kiŋ(ba)  | ka(we:ŋ)             | ku(buŋ)            |
| 'mouth' <sup>23</sup>  | *-wa                | -aw              | -wa       | -ua     | -wa(r)              | -va      | -(ar)fa  |         |         | -wa   | -wa:     | -wa                  | -wa                |
| 'name'                 | *-en(i,u)           |                  | -einu     | -en     | -in:u               | -ene     | -ni      | -ne?    | -enei   | -ne   | -nei     | -ni                  |                    |
| 'new'                  | *siba               | ɸsib             | sava(?a)  | sib     | sab:a <sup>24</sup> | hiba     | haba(r)  | həba    | saba    | tifa  | supa(ka) | tiŋpa                | təpa               |
| 'nose'                 | *-min               |                  |           | -min    | ɸ-mi                | -miŋ     | -miŋ     | -muin   | -min    |       |          | -miŋi                | -muŋi              |
| 'one'                  | *nuk                | nuk              | nuku      | nuk     | nuku                | nu       | nu       | nuk     | nuku    | nuku  | nok      |                      | no                 |
| 'pierce' <sup>25</sup> | *tapai              | tap              | tapa      | tap     | tap:a(t)            | tapa     | tap:a(t) | tapə(n) | tapə(n) | tapai | tapei    | tapea                | ta                 |
| 'pig'                  | *baj                | bai              | bei       | bei     | bai                 | bε       | boi      | be?     | bei     | fe    | pe       | pi                   | pei                |
| 'plant' (v.)           | *mudin              | midan            | mudi      | mudun   | mid:iŋ              | mud:iŋ   | midin    | midin   | medi    | murui | mit      | mad:iŋ               | mədi               |
| 'rat'                  | *dur                | dur              | dur       | dur     | di                  | duru     | ɸdur     | dur     | dur     | dur   | tui      | daru                 | dur(ki)            |
| 'rattan'               |                     | liag             | liag      | le:g    |                     | ?ilia    | le       |         | le      |       |          |                      |                    |
| 'recline'              | *tia                | ti:?             | ta?a      | te      | ti?an               | tia      | ta:      | ta:     | ta:     | ta:   | ta:      | -te                  | taj                |
| 'right'                |                     | jidan            | jediŋ     | jad:iŋ  |                     |          |          |         |         |       |          |                      |                    |

2 The internal history of the Alor-Pantar language family

| Gloss                | pAP                      | Tw                 | Nd                  | Ke                  | WP                | Bl       | Ad      | Kl      | Ki      | Ab                | Km                 | Sw        | We       |
|----------------------|--------------------------|--------------------|---------------------|---------------------|-------------------|----------|---------|---------|---------|-------------------|--------------------|-----------|----------|
| 'ripe'               | *tena                    |                    |                     | ten                 | tena              | ten      | etɛn    | tain    | iten    | itina             |                    |           |          |
| 'roof'               | *wai                     | wai                | waja                |                     | wai               | vai      | fa      | wei     | wai     | wai               | itɛn <sup>27</sup> |           |          |
| 'rotten'             | *mVn                     | mʉn                | -mini               | mino                | min(isa)          | min(isa) | ʃmul    | mʉin    |         | -mun              |                    |           |          |
| 'saltwater'          | *tam                     | ʃta?               | ta                  | tam                 | tawa              | taŋ      | taŋ     | tan     | tama    | tama              | tama               | tama      | tama?    |
| 'scorpion'           | *pVr                     | par                | par                 | par                 | ʃpar              | ʃpɛl     | pail    | par     | pe-i    | pe-i              | ʃfal               |           | per(buk) |
| 'search'             |                          |                    | lafi                | rap                 | rapin             | rapin    | lap     | -rap    |         |                   |                    |           |          |
| 'shark'              | *sib(a.i)r <sup>28</sup> | sipar              | sifi                | sibar               | sib-u             | hibir    | ʃtabei  | sobor   |         |                   |                    |           |          |
| 'short'              | *tukV                    | tuk                | tuku                | tuk                 | tuka(n)           | tuka(n)  | toʔa(n) | tuk     | tuk     | tuk               | tuk <sup>29</sup>  | tuku(da)  | tuk      |
| 'sibling (older)'    | *nan(a)                  |                    | -nan                | da:ro <sup>32</sup> | dali              | dar      | dal     | dar     | nana    | nana              |                    | -na:na    | -nan     |
| 'sing'               | *dar(a)                  | da:r               | da:la <sup>31</sup> |                     |                   |          |         |         | dar     | jaʃ <sup>33</sup> |                    | dara      | dəra     |
| 'sit'                | *mis                     | mis                | mi                  | mis-                | mis(in)           | mihi     | mih     | mih     | misa    | mit               | ʃnih               | miti      | amit     |
| 'six'                | *talam                   | ʃtia:m             | ʃtiama              | ʃtiam               | taliŋ             | taliŋ    | talaŋ   | təlan   | talama  | tala:ma           | ta:ma              |           |          |
| 'sky'                |                          | bulan              |                     | buluŋ               | ʃburan            | ʃburan   |         |         |         |                   |                    |           |          |
| 'slippery'           | *dul(a)                  |                    |                     | duj-                | ʃduba             | dula     | dul     | du:l    | dula    | rula              | tula(ka)           | dalo:(ka) | dol(ok)  |
| 'smoke'              | *bunaq                   | bun                | bun                 | banax               | bun:a             | benaka   | bano?   | -bon    | bonok   | puna              | puna               | punaka    | punak    |
| 'spear'              | *qaba(k)                 | ʃqab               | ʃqaba               | xabi                | kab:i             | ʔaba     | ʔaba    | kəbak   | kəbak   | kafak             | kapa               |           |          |
| 'spit'               | *purVn                   | puran              | (mali)plum          | puran               | purun             | tahi     | pui     | pəruin  | purin   | puina             | (su)pui            |           |          |
| 'stand'              | *tas                     | tas                | tasi                | tas-                | ʃtib              | ʃtib     | toh     | (mə)teh | (na)tet |                   |                    |           | -tati    |
| 'star'               | *jibV                    | jiʃ                | ifa(xoja)           | ʃip(alaq)           | hib:i             | ʃtid     | ib(in)  | ʔib     | ib(ra)  |                   |                    |           |          |
| 'stomach'            | *tok                     | ʃ-to?              | ʃ-toʔo              | -toki               | -tow              | -tow     | -to?    |         |         |                   | -tok               |           | -toko    |
| 'stone'              | *war                     | war                | wala                | war                 | war               | war      | ʃwi     | wɔr     | wor     | wi                | wor                | wara      | wor      |
| 'sugarcane'          | *huba                    | ʃhub <sup>34</sup> | u:fa                | u:b                 | habua             | ub       | ʃsob    | aba     | u:b     |                   |                    |           | upa      |
| 'sun'                | *wadi                    | war (get)          | weri                | ʃwer                | war <sup>35</sup> | wed      | fed     |         | ʃber    | war               | wati               | wadi      | widi     |
| 'tail'               | *ora                     | -or                | ola                 | -or                 | -ora              | -ora     | ol      | -or     | -or     | -wai              | (w)ui              | (w)ora    | (w)ori   |
| 'tooth'              | *uasin                   | usan               | usit                | uasit               | wasit             | ʃ-veit   | ʃhit    | -weh    | -wes    | -weti             | -weh               | ʃ-wa      | ʃ-wesi   |
| 'tens' <sup>36</sup> | *qar-                    | qa:r-              | qa-                 | xar                 | ke-               | ʔar-     | ʃ?er-   | kar-    | kar-    | ʃkar-             |                    |           |          |

| Gloss                | pAP                    | Tw      | Nd      | Ke               | WP    | Bl      | Ad               | Kl                  | Ki     | Ab      | Km                 | Sw         | We     |
|----------------------|------------------------|---------|---------|------------------|-------|---------|------------------|---------------------|--------|---------|--------------------|------------|--------|
| 'thatch'             | *amen                  | man     | man     | man              |       | meniy   | men              | enɛ:m <sup>37</sup> | amen   | ameny   |                    | ameny      | amen   |
| 'thick'              | *dumV                  | ɕtuʔum  |         |                  | dum:a |         |                  |                     |        |         |                    | dumu       | dum    |
| 'throw'              | *oda                   |         |         | od               |       | oda     | od               | o:d                 | or     | ɕwot    | woia <sup>38</sup> |            |        |
| 'tongue'             | *-lebur                | -livi   | lefu    | -leb             | -lebu | ɕ-lebul | -libu(ŋ)         | -leb                | liber  | lifi    | ɕ-opui             | -li(m)puru | ɕjebur |
| 'tree'               | *tei                   | tei     | tei     | tei              |       | te      | ti               | (ɛ)teʔ              | (a)tei | (ba)taa |                    |            |        |
| 'two'                | *araq                  | raq     | ɕraq    | rax              | alaku | ɕaru    | ab               | orok                | oruku  | ajoku   | ɕok                | ɕjaku      | ɕjoku  |
| 'vagina'             | *-ar                   | -ar     |         | abən             | habən | -ar     | -al              | -ar                 | -ar    | -oi     | -ai                | ɕ-la       |        |
| 'village'            | *haban                 | haɕan   | aɕan    | abən             | habən | abən    | ban              | ɛben                | aban   | aɕen    | -tan               |            | -teŋ   |
| 'wake s.o.'          | *-ten                  |         | -tani   | -ten             |       |         |                  | -ten                |        |         |                    |            |        |
| 'walk <sub>1</sub> ' | *lam(ar) <sup>39</sup> | lam(an) |         | ɕamar            | lama  | ɕlamal  | lamɛ             | lam                 |        |         |                    |            |        |
| 'walk <sub>2</sub> ' |                        |         |         |                  |       |         |                  |                     |        |         |                    |            |        |
| 'water'              | *jira                  | jir     | jila    | ir               | hila  | dzar    | sei              | ara:                | er     | lol     | lo:                | lo:la      | lailol |
| 'wave'               | *bob                   | boɕ     | bova    | bo:b             | ɕbo   | bob     | bob              | bo:b                | fɔi    |         |                    | iria       | ira    |
| 'white'              |                        | miaq    | miaqa   | miex             | miaka | madzaka |                  |                     |        |         |                    |            |        |
| 'wind' <sup>41</sup> |                        |         |         |                  |       |         | hamoi            |                     |        | timoi   | sumui              | tamuro     |        |
| 'wound'              |                        | bat     | bata    | ibat             |       | bata    | ɕbah             | ɕabad               | bata   |         |                    |            |        |
| 'yawn'               | *hagur                 | haɕar   |         | agur             |       |         |                  |                     | ɕagu   | ahau    |                    |            |        |
| 'yellow'             | *bagori <sup>42</sup>  | baɕari  | ɕbaxori | bagari           | bug-u | boʔori  | ɕbaʔoil          | bəgor               | bagura |         |                    |            |        |
| 1PL.INCL             | *pi-                   | pi-     | pi-     | pi-              | pi-   | pi-     | pi-              | pi-                 | pi-    | pi-     | na-                | pi-        | ne-    |
| 1SG                  | *na-                   | na-     | na-     | na-              | na-   | na-     | na-              | na-                 | na-    | na-     | na-                | na-        | a-     |
| 2SG                  | *ha-                   | ha-     |         | ga-              | gai-  | ʔe-     | ʔe <sup>43</sup> | ge-                 | he-    | ho-     | ge-                | ge-        | a-     |
| 3GEN                 | *ge-                   |         |         |                  |       | ʔo-     | ʔo-              | go-                 | ho-    | ho-     | wo-                |            |        |
| 3LOC                 |                        |         |         |                  |       |         |                  |                     |        |         |                    |            |        |
| 3PL                  | *gi-                   | gi-     | ga-     | gV <sup>44</sup> | ga-   | ʔi-     | ʔa-              | g-                  | ga-    | ha-     | ga-                | gi-        | gi-    |
| 3SG                  | *ga-                   | ga-     | ga-     | gV <sup>44</sup> | ga-   | ʔa-     | ʔa-              | g-                  | ga-    | ha-     | ga-                | ga-        | gV-    |

## Notes to tables

<sup>1</sup> A form \*balin ‘axe’ was reconstructed to pAP in Holton et al. (2012), but we now recognize that this is an Austronesian loan, probably from Alorese *baling*.

<sup>2</sup> This form has metathesized.

<sup>3</sup> This form has metathesized.

<sup>4</sup> Denotes ‘chicken’.

<sup>5</sup> 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 neighbors of the Alor-Pantar languages.

<sup>6</sup> Denotes ‘traditional dance’.

<sup>7</sup> Denotes ‘bow, bend’.

<sup>8</sup> This form was not reconstructed to pAP in Holton et al. (2012) because it is not attested in Alor languages. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>9</sup> Denotes ‘not quite dry’.

<sup>10</sup> Denotes ‘eat’ in Tw, Nd, WP, Ab, Km, ‘eat/drink’ in Ke, Bl, Sw, and ‘drink’ in Ad, Kl, Ki, We.

<sup>11</sup> This form exhibits metathesis.

<sup>12</sup> This form was not reconstructed to pAP in Holton et al. (2012) because of its limited distribution. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>13</sup> Denotes ‘burn, of land’.

<sup>14</sup> Denotes ‘clothing louse’, with metathesis.

<sup>15</sup> Note similarity with proto-Austronesian \*isi? ‘contents’, indicating that this may be a loan.

<sup>16</sup> Denotes ‘meat’.

<sup>17</sup> This form was not reconstructed to pAP in Holton et al. (2012) because it is not attested in any Pantar language. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>18</sup> Reflexes of \*p{i,u}nV typically encompass the meanings ‘hold’ and ‘grab’ with the difference depending on the prefixation of the verb.

<sup>19</sup> Sw has *wuni* ‘hold’ and *puni* ‘hit’.

<sup>20</sup> We has *woiŋ* ‘hold’ and *poiŋ* ‘hit’.

<sup>21</sup> This form was not reconstructed to pAP in Holton et al. (2012) due to its limited distribution. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>22</sup> Denotes ‘maggot’.

<sup>23</sup> This form is generally part of a compound when meaning ‘chin’. It seems to have historically meant ‘mouth’. It is retained with that meaning in Ke, Ki, Ab, and Km. In Tw, Nd, WP, Bl, Ad, Sw, and We, the form is only retained as part of a compound meaning ‘chin’.

<sup>24</sup> Denotes ‘new sprout’.

<sup>25</sup> Reflexes of \*tapai encompass the meanings ‘pierce’, ‘stab’ ‘sew’, ‘plant in the ground’, and ‘pound rice’.

<sup>26</sup> Denotes ‘on top’.

<sup>27</sup> Denotes ‘thatch’.

<sup>28</sup> This form was not reconstructed to pAP in Holton et al. (2012) because it is not attested in the eastern languages. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>29</sup> Denotes ‘piece, chunk’.

<sup>30</sup> Denotes ‘short piece, cutting’.

<sup>31</sup> Denotes ‘dance’.

<sup>32</sup> Denotes ‘dance’.

<sup>33</sup> This form has lost the initial syllable.

<sup>34</sup> Denotes ‘sweet’.

<sup>35</sup> Denotes ‘shine, burn’ (cf. *was* ‘sun’).

<sup>36</sup> Given that the Abui reflex is irregular, strictly speaking this set does not meet the distributional criteria for reconstruction, since there is no regular reflex in Eastern Alor.

<sup>37</sup> This form has metathesized.

<sup>38</sup> Denotes ‘beat, strike (drum)’.

<sup>39</sup> This form was not reconstructed to pAP in Holton et al. (2012) because it is not attested in the eastern languages. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>40</sup> Tw *laman* ‘follow, walk along (e.g. a path)’. WP *lama* shares this sense and is likely a borrowing from Tw, which explains the lack of gemination in the WP form.

<sup>41</sup> This may be an Austronesian loan. Note proto-Malayo-Polynesian \*timuR ‘south-east monsoon’ (Blust & Trussel 2010).

<sup>42</sup> This form was not reconstructed to pAP in Holton et al. (2012) because it is not attested in any eastern language. However, based on its presence in Timor languages (see Chapter 3), we now reconstruct it to pAP.

<sup>43</sup> In Adang *?e* has been restricted to marking possessors in contrastive focus.

<sup>44</sup> Prefix vowel harmonizes with stem vowel.

## Acknowledgements

This paper draws on the work of Holton et al. (2012) and Robinson & Holton (2012). We are particularly grateful to our co-authors of the former paper for sharing data and insights regarding the reconstruction of Proto-Alor-Pantar. Those authors are not to be blamed for errors of fact or interpretation resulting from the revisions found in the current chapter.

## Abbreviations

|      |                        |
|------|------------------------|
| 1    | 1st person             |
| 2    | 2nd person             |
| 3    | 3rd person             |
| AB   | Abui                   |
| AD   | Adang                  |
| AN   | Austronesian languages |
| AP   | Alor-Pantar            |
| BL   | Blagar                 |
| ENV  | environment            |
| GEN  | Genitive               |
| INCL | Inclusive              |
| KE   | Kaera                  |
| KI   | Kui                    |
| KL   | Klon                   |
| KM   | Kamang                 |
| LOC  | Locative               |
| ND   | Nedebang               |
| PAP  | proto-Alor-Pantar      |
| PL   | Plural                 |
| SG   | Singular               |
| SW   | Sawila                 |
| TAP  | Timor-Alor-Pantar      |
| Tw   | Teiwa                  |
| v    | verb                   |
| V    | vowel                  |
| WE   | Wersing                |
| WP   | Western Pantar         |

## References

- Anonymous. 1914. De eilanden Alor en Pantar, residentie Timor en Onderhoorigheden. *Tijdschrift van het Koninklijk Nederlandsch Aardrijkskundig Genootschap* 31. 70–102.
- Baird, Louise. 2008. *A grammar of Klon: a non-Austronesian language of Alor, Indonesia*. Canberra: Pacific Linguistics.
- Blust, Robert Andrew & Stephen Trussel. 2010. *Austronesian Comparative Dictionary, web edition*. <http://www.trussel2.com/acd/>.
- de Josselin de Jong, Jan Petrus Benjamin. 1937. *Studies in Indonesian culture I: Oirata, a Timorese settlement on Kisar*. Amsterdam: Noord-Hollandsche Uitgevers-Maatschappij.
- Drummond, Alexei J., M. A. Suchard, D. Xie & A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29(8). 1969–1973.
- Du Bois, Cora. 1944. *The people of Alor: a social-psychological study of an East Indian island*. Minneapolis: University of Minnesota Press.
- Dunn, Michael, Niclas Burenhult, Nicole Kruspe, Sylvia Tufvesson & Neele Becker. 2011. Aslian linguistic prehistory: a case study in computational phylogenetics. *Diachronica* 28(3). 291–323.
- Hajek, John. 2010. Towards a phonological overview of the vowel and consonant systems of East Nusantara. In Michael C. Ewing & Marian Klamer (eds.), *Typological and areal analyses: Contributions from East Nusantara*, 25–46. Canberra: Pacific Linguistics.
- Holland, B. R., K. T. Huber, A. Dress & V. Moulton. 2002.  $\Delta$  Plots: a tool for analyzing phylogenetic distance data. *Molecular Biology and Evolution* 19(12). 2051–2059.
- Holton, Gary, Marian Klamer, František Kratochvíl, Laura C. Robinson & Antoinette Schapper. 2012. The historical relations of the Papuan languages of Alor and Pantar. *Oceanic Linguistics* 51(1). 86–122.
- Huson, D. H. & D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2). 254–267.
- Klamer, Marian. 2010. *A grammar of Teiwa* (Mouton Grammar Library 49). Berlin: Mouton de Gruyter.
- Klamer, Marian. 2011. *A short grammar of Alorese (Austronesian)* (Languages of the World/Materials 486). München: Lincom.
- Klamer, Marian. 2012. Papuan-Austronesian language contact: Alorese from an areal perspective. In Nicholas Evans & Marian Klamer (eds.), *Melanesian lan-*



## 2 The internal history of the Alor-Pantar language family

- languages on the edge of Asia: challenges for the 21st century*, vol. 5 (Language Documentation & Conservation Special Publication), 72–108. Honolulu: University of Hawaii Press.
- Kratochvíl, František, Sebastian Fedden, Gary Holton, Marian Klamer, Laura C. Robinson & Antoinette Schapper. 2011. *Pronominal systems in Alor-Pantar languages: Development and diachronic stability*. Paper presented at the International Conference on Historical Linguistics. Osaka, July 28.
- Lewis, Paul M. (ed.). 2009. *Ethnologue: languages of the world*. 16th edn. Dallas: SIL International. <http://www.ethnologue.com/web.asp>.
- Lichtenberk, František. 1998. Did speakers of Proto Oceanic chew betel? *Journal of the Polynesian Society* 107. 335–363.
- Maddieson, Ian. 2005. Consonant inventories. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 10–13. Oxford: Oxford University Press.
- Mona, Stefano, Katharina E. Grunz, Silke Brauer, Brigitte Pakendorf, Loredana Castri, Herawati Sudoyo, Sangkot Marzuki, Robert H. Barnes, Jörg Schmidtke, Mark Stoneking & Manfred Kayser. 2009. Genetic admixture history of Eastern Indonesia as revealed by y-chromosome and mitochondrial DNA analysis. *Molecular Biology and Evolution* 26(8). 1865–1877.
- Nicolspeyer, Martha Margaretha. 1940. *De sociale structuur van een Aloreesche bevolkingsgroep*. Rijswijk: Rijksuniversiteit te Leiden PhD thesis.
- Pagel, Mark. 2009. Human language as a culturally transmitted replicator. *Nature Reviews Genetics* 10(6). 405–415.
- Pawley, Andrew K. 2005. The chequered career of the Trans New Guinea hypothesis: recent research and its implications. In Andrew K. Pawley, Robert Attenborough, Jack Golson & Robin Hide (eds.), *Papuan pasts: cultural, linguistic and biological history of the Papuan-speaking peoples*, 67–108. Canberra: Pacific Linguistics.
- Rambaut, Andrew & Alexei J. Drummond. 2007. *Tracer, vers. 1.5*. <http://beast.bio.ed.ac.uk/Tracer>.
- Robinson, Laura C. & John Haan. 2014. Adang. In Antoinette Schapper (ed.), *Papuan languages of Timor, Alor and Pantar: Sketch grammars*, vol. 1, 221–284. Berlin: Mouton de Gruyter.
- Robinson, Laura C. & Gary Holton. 2012. Internal classification of the Alor-Pantar language family using computational methods applied to the lexicon. *Language Dynamics and Change* 2(2). 123–149.
- Ronquist, F. & J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(1572-1574).

- Schapper, Antoinette (ed.). 2014. *Papuan languages of Timor, Alor and Pantar: Sketch grammars*. Vol. 1. Berlin: Mouton de Gruyter.
- Schapper, Antoinette, Juliette Huber & Aone van Engelenhoven. 2012. The historical relation of the Papuan languages of Timor and Kisar. In Harald Hammarström & Wilco van den Heuvel (eds.), *History, contact and classification of Papuan languages* (Language and Linguistics in Melanesia Special Issue 2012, Part I), 194–242. Port Moresby: Linguistic Society of Papua New Guinea.
- Schapper, Antoinette, Juliette Huber & Aone van Engelenhoven. this volume. The relatedness of Timor-Kisar and Alor-Pantar languages: A preliminary demonstration. In Marian Klamer (ed.), *The Alor-Pantar languages*, 99–154. Berlin: Language Science Press.
- Spriggs, Matthew. 2011. Archaeology and the Austronesian expansion: where are we now? *Antiquity* 85. 510–528.
- Stokhof, W. A. L. 1975. *Preliminary notes on the Alor and Pantar languages (East Indonesia)* (Pacific Linguistics: Series B 43). Canberra: Australian National University.
- Wichmann, Søren, Eric W. Holman, Taraka Rama & Robert S. Walker. 2011. Correlates of reticulation in linguistic phylogenies. *Language Dynamics and Change* (1). 205–40.
- Wurm, Stephen A. 1982. *The Papuan languages of Oceania* (Ars Linguistica 7). Tübingen: Gunter Narr.
- Yen, Douglas E. 1977. Hoabinhian horticulture? The evidence and the questions from northwest Thailand. In Jim Allen, Jack Golson & Rhys Jones (eds.), *Sunda and Sahul: Prehistoric studies in Southeast Asia, Melanesia and Australia*, 567–599. London: Academic Press.