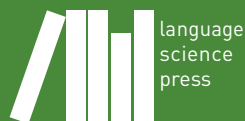


# The Alor-Pantar languages

History and typology

Marian Klamer (ed.)

Studies in Diversity Linguistics, No ??



## Studies in Diversity Linguistics

Chief Editor: Martin Haspelmath

Consulting Editors: Fernando Zúñiga, Peter Arkadiev, Ruth Singer, Pilar Valenzuela

In this series:

1. Handschuh, Corinna. A typology of marked-S languages
2. Rießler, Michael. Adjective attribution
3. Klammer, Marian (ed.). The Alor-Pantar languages: History and typology
4. Berghäll, Liisa. A grammar of Mauwake (Papua New Guinea)
5. Wilbur, Joshua. A grammar of Pite Saami

# The Alor-Pantar languages

History and typology

Marian Klamer (ed.)

Language Science Press  
Berlin

Language Science Press  
Habelschwerdter Allee 45  
14195 Berlin, Germany

[langsci-press.org](http://langsci-press.org)

This title can be downloaded at:

<http://langsci-press.org/catalog/book/22>

© 2014, Marian Klammer (ed.)

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/>

ISBN: 978-3-944675-48-0

Cover and concept of design: Ulrike Harbort  
Typesetting: Sebastian Nordhoff; Timm Lichte  
Proofreading: Benedikt Singpiel

Storage and cataloguing done by FU Berlin



Language Science Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables and other factual information given in this work are correct at the time of first publication but Language Science Press does not guarantee the accuracy of such information thereafter.

This volume is dedicated to the memory of Mr. Anderias Malaikosa (1964-2011), whose love for the peoples and languages of his native Alor was an inspiration to linguists both within Alor and abroad.



# Contents

<b>1 Numeral systems in the Alor-Pantar languages</b> Antoinette Schapper & Marian Klammer	<b>1</b>
<b>Language Index</b>	<b>51</b>
<b>Name Index</b>	<b>56</b>
<b>Subject Index</b>	<b>56</b>





## Chapter 1

# Numeral systems in the Alor-Pantar languages

Antoinette Schapper & Marian Klammer

This chapter presents an in-depth analysis of numeral forms and systems in the Alor-Pantar (AP) languages. The AP family reflects a typologically rare combination of mono-morphemic ‘six’ with quinary forms for numerals ‘seven’ to ‘nine’, a pattern which we reconstruct to go back to proto-AP. We focus on the structure of cardinal numerals, highlighting the diversity of the numeral systems involved. We reconstruct numeral forms to different levels of the AP family, and argue that AP numeral systems have been complicated at different stages by reorganisations of patterns of numeral formation and by borrowings. This has led to patchwork numeral systems in the modern languages, incorporating to different extents: (i) quaternary, quinary and decimal bases; (ii) additive, subtractive and multiplicative procedures, and; (iii) non-numeral lexemes such as ‘single’ and ‘take away’. Complementing the historical reconstruction with an areal perspective, we compare the numerals in the AP family with those of the Austronesian languages in their immediate vicinity and show that contact-induced borrowing of forms and structures has affected numeral paradigms in both AP languages and their Austronesian neighbors.

## 1 Introduction

Numerals and numeral systems have long been of typological and historical interest to linguists. Papuan languages are best known in the typological literature on numerals for having body-part tally systems and, to a lesser extent, restricted numeral systems which have no cyclically recurring base (Laycock 1975, Lean 1992, Comrie 2005a). Papuan languages are also typologically interesting for the fact that they often make use of bases of other than the cross-linguistically most



frequent decimal and vigesimal bases, such as quinary (Lean 1992) and senary bases (Donohue 2008, Evans 2009).

In this chapter we present an in-depth analysis of numeral forms and systems in the Alor-Pantar (AP) languages. Typologically, the family reflects a rare combination of mono-morphemic ‘six’ with quinary forms for numerals ‘seven’ to ‘nine’, a pattern which we reconstruct back to proto-AP. We focus on the structure of cardinal numerals, highlighting the diversity of the numeral systems involved. We reconstruct numeral forms to different levels of the AP family, and argue that AP numeral systems have been complicated at different stages by reorganisations of patterns of numeral formation and by borrowings. This has led to patchwork numeral systems in the modern languages, incorporating to different extents: (i) quaternary, quinary and decimal bases; (ii) additive, subtractive and multiplicative procedures, and; (iii) non-numeral lexemes such as ‘single’ and ‘take away’. We complement the genealogical perspective with an areal one, comparing the numeral systems of the AP languages with those of the Austronesian languages in their immediate vicinity to study if and how contact has affected the numeral paradigms.

This chapter centres on numeral data from 19 Alor-Pantar language varieties spanning east to west across the AP archipelago, presented collectively in Appendix A.1. As a motivated phonemic orthography is yet lacking for many of the varieties in our sample, all the data is presented in a broad IPA transcription. The fieldworkers who collected the data are recognised in the ‘Sources’ section at the end of the chapter.

We begin with an overview of the terminology used throughout this chapter in § 2 and a brief note on sound changes in numeral compounds in § 3. We then describe how cardinal numerals are constructed across the AP languages: ‘one’ to ‘five’ are discussed in § 4, ‘six’ to ‘nine’ in § 5, and numerals ‘ten’ and above in § 6. § 7 looks at the AP numeral systems in typological and areal perspective, while § 8 summarises our findings.

## 2 Terminological preliminaries

Numerals are ‘spoken normed expressions that are used to denote the exact number of objects for an open class of objects in an open class of social situations with the whole speech community in question’ (Hammarström 2010: 11). A numeral system is thus the arrangement of individual numeral expressions together in a language.

Numeral systems typically make use of a base to construct their numeral ex-

pressions.<sup>1</sup> A “base” in a numeral system is a numerical value  $n$  which is used repeatedly in numeral expressions thus:  $xn \pm/xy$ , that is, numeral  $x$  is multiplied by the base  $n$  plus, minus or multiplied by numeral  $y$  (Comrie 2005b, Hammarström 2010: 15).<sup>2</sup> Many languages have multiple bases. For instance, Dutch numerals have five different bases: *tien* ‘10’, *honderd* ‘100’, *duizend* ‘1000’, *miljoen* ‘100,000’, *miljard* ‘1,000,000’. These bases are all powers of ten ( $10$ ,  $10^2$ ,  $10^3$ ,  $10^6$ ,  $10^9$ ). However, the higher powers are not typically considered important in defining a numeral system type; the lowest base gives its name to the whole system, that is, Dutch would be characterised as a “decimal” or “base-10” numeral system.

In this chapter we deal with several “mixed numeral systems”. We define a “mixed numeral system” as a numeral system in which there are multiple bases that are *not* simply powers of the lowest base. So, we do not consider Dutch to have a mixed numeral system, since all its higher bases are powers of its lowest base, *tien* ‘10’. By contrast, a language such as Ilongot (Table 1) would be considered to have a mixed quinary-decimal system because: (i) it uses a quinary base to form numerals ‘six’ to ‘nine’, and (ii) a decimal base to form numerals ‘ten’ and higher. ‘Ten’ is not a power of ‘five’ and therefore the language can be considered to “mix” numeral bases.

It is important to note that isolated cases of a particular mathematical procedure being used in the formation of a numeral do not constitute an instance of another ‘base’ in a numeral system. For instance, Ujir (Table 1) forms ‘seven’ by means of the addition of ‘six’ and ‘one’. Yet ‘six’ is not a base in Ujir, since there are no other numerals in the language formed with additions involving ‘six’. Similarly, ‘two’ and ‘four’ are not bases in Ujir, because neither is used recursively in forming numerals. The formation of ‘eight’ through the multiplication of ‘two’ and ‘four’ is a procedure limited to ‘eight’.

In this chapter, we are concerned with the internal composition of cardinal numerals, that is, if and how they are made up out of other numeral expressions. We call a monomorphemic cardinal a “simplex numeral”, and one that is composed of more than one numeral expressions a “complex numeral”. To describe (i) the arithmetic relation between component elements in a complex numeral, and (ii) the role of component elements in arithmetic operations, the following

<sup>1</sup> Notable exceptions, i.e., numeral systems without bases, are the body-tally systems mentioned above, and the languages discussed in Hammarström (2010:17-22).

<sup>2</sup> We do not adopt the notion of ‘base’ of Greenberg (1978) where ‘base’ is defined as a serialised multiplicand upon which the recursive structure of *all* higher complex numerals is constructed. That is, even where a language has for instance a small sequence of numerals formed on a non-decimal pattern (e.g., ‘5 2’ for ‘seven’, ‘5 3’ for ‘eight’, and ‘5 4’ for ‘nine’), if ‘10’ is the higher, more productive base, then the language is classed as having a decimal system only.

Table 1: Examples illustrating the notion of “base”

Ilongot			Ujir		
Austronesian			Austronesian		
Philippines			Indonesia		
	Analysis	Expression		Analysis	Expression
1	1	<i>sit</i>	1		<i>set</i>
2	2	<i>dewa</i>	2		<i>rua</i>
3	3	<i>teyo</i>	3		<i>lati</i>
4	4	<i>opat</i>	4		<i>ka</i>
5	5	<i>tambian</i>	5		<i>lima</i>
6	5 + 1	<i>tambianno sit</i>	6		<i>dubu</i>
7	5 + 2	<i>tambianno dewa</i>	6 + 1		<i>dubusam</i>
8	5 + 3	<i>tambianno teyo</i>	4 x 2		<i>karua</i>
9	5 + 4	<i>tambianno opat</i>	9		<i>tera</i>
10	10	<i>tampo</i>	10		<i>uisia</i>
11	10 + 1	<i>tampo no sit</i>	10 + 1		<i>uisia ma set</i>
15	10 + 5	<i>tampo no tambian</i>	10 + 5		<i>uisia ma lima</i>
20	2 x 10	<i>dowampo</i>	2 x 10		<i>uirua</i>
21	2 x 10 + 1	<i>dowampo no sit</i>	2 x 10 + 1		<i>uirua ma set</i>
25	2 x 10 + 5	<i>dowampo no tambian</i>	2 x 10 + 5		<i>uirua ma lima</i>
30	3 x 10	<i>teyompo</i>	3 x 10		<i>uilati</i>
<b>BASES</b>	5-10			10	

terms are used:

- “additive numeral”: a numeral where the relation between components parts of a complex numeral is one of addition. The component parts are “augend” and “addend”. So, for example, in the equation  $5 + 2 = 7$ , the augend is 5 and the addend is 2.
- “subtractive numeral”: a numeral where the relation between component parts of a complex numeral is one of subtraction. The component parts are “subtrahend” and “minuend”. So, for example, in the equation  $10 - 2 = 8$ , the subtrahend is 2 and the minuend is 10.
- “multiplicative numeral”: a numeral where the relation between components parts of a complex numeral is one of multiplication. The component parts are “multiplier” and “multiplicand”. So, for example, in the equation  $3 \times 2 = 6$ , the multiplier is 3 and the multiplicand is 2.

Throughout this chapter we rely on the definitions made in this section, and the reader is referred to this section for clarification of terminology.

### 3 A brief note on sound changes and numerals

In this chapter we posit reconstructions of numerals to proto-Alor-Pantar (pAP) and several lower subgroups within the AP group. Many of the sound correspondences on which these reconstructions are based are part of regular correspondence sets discussed in Holton et al. (2012) and Holton & Robinson (this volume).

However, the history of numerals also involves formal changes which cannot be couched in terms of regular sound correspondences. Many irregular changes observed in numerals arise from members of compounds fusing together over time. In the history of AP numerals, two types of change are associated with the compounding process: (i) segmental reduction in the members of a compound, (ii) dissimilation of segments across members of a compound.

Examples of segmental reduction in numeral compounds are widespread in AP languages. For instance, in the Atoitaa dialect of Kamang, numerals ‘seven’ to ‘nine’ are formed with *iwesij* ‘five’ followed by a numeral ‘one’ to ‘four’. This is illustrated for ‘six’ in (1). In forming the compound, the medial syllable of ‘five’, /we/, is lost due to a shift in stress to the penultimate syllable. Unreduced forms involving two distinct phonological forms are only produced by speakers when explaining numeral formation and have not been observed in naturalistic speech,

indicating that the reduced form is already well incorporated into speakers' lexicons.

- (1) Variation in the realisation of Kamang (Atoitaa) 'six'
  - a. *iwesij nok* [i'wesij 'nok] (careful speech)  
five one  
'six'
  - b. *isijnok* [i'sijnok] (normal speech)  
five.one  
'six'

Similarly, in Sawila we find that 'six' can be realised both in unreduced form as two distinct numerals (*jo:tiŋ* 'five' [plus] *suna* 'one') and in reduced form as set out in (2).

- (2) Variation in the realisation of Sawila 'six'
  - a. *jo:tiŋ suna* (careful speech)  
five one  
'six'
  - b. *jo:tsuna* (normal speech)  
five.one  
'six'

Dissimilation of segments across members of a compound is also found. An example is Klon *tidorok* 'eight', a form which must have involved consonant dissimilation of the protoform \**turarok* (see Table 5) and a hypothetical intermediate form like \*\**tudarok* (§ 5.2.2).

In short, the reconstruction of numerals must take into account regular sound changes as well as irregular changes in the members of compounds.

## 4 Numerals 'one' to 'five'

The numerals 'one' to 'five' are for the most part simple mono-morphemic words in Alor- Pantar. Table 2 presents an overview with the reconstructions to proto-Alor-Pantar (pAP).<sup>3</sup> The Proto-AP numerals 'one' to 'five' have been retained in

<sup>3</sup> Not all elements of the reconstructed forms as they are given here are motivated in this chapter; see the reconstructed sound changes reported on in Holton et al. (2012) and Holton & Robinson (this volume).

most of its descendants. Only in eastern Alor have numerals in this range been innovated.

A non-etymological initial /a/ is present on Western Pantar ‘one’ and ‘four’ and Reta ‘one’. This development is apparently due to analogy with the numerals ‘two’ and possibly ‘three’. Such analogical adjustments in numeral forms, sometimes referred to as ‘onset runs’ (Matisoff 1995), are cross-linguistically relatively common.<sup>4</sup> The prothetic /a/ is also found on Western Pantar ‘thousand’ which can be realised as either *ribu* or *aribu*, an Austronesian loan.

The etymologies of the AP numerals ‘two’ and ‘five’ have been the subject of some speculation and typological interest. In his early comparative study on Alor-Pantar languages, Stokhof (1975: 21) makes two observations about these AP numerals which are not supported by our data. First, his claim that AP languages frequently use the root ‘tooth’ to express ‘five’ is not supported by more recent historical work on the family, which reconstructs ‘five’ as pAP \*jiwesin, and ‘tooth’ as \*-uas(in) (Holton & Robinson this volume). It should also be noted that no known cognitive link exists between ‘five’ and ‘tooth’, unlike the link between ‘five’ and ‘hand’ (Majewicz 1981, 1984, Heine 1997). Second, contra Stokhof, pAP \*(a)tiga ‘three’ is not a loan word from Malay, despite the similarity with Malay *tiga* ‘three’. Whilst there is evidence of Austronesian influence in pAP,<sup>7</sup> there is no evidence of influence from Malay. Malay first arrived in the Alor-Pantar region in colonial times,<sup>8</sup> thousands of years after the likely break-up of

<sup>4</sup> For example, in the Austronesian language Thao (Taiwan), initial /s/ in \*susha ‘two’ was replaced by /t/ (*tusha*) in analogy to the onsets of *ta* ‘one’ and *turu* ‘three’ (Blust 2009:274). The initial /d/ on ‘nine’ in Slavonic languages (e.g., Russian *dévjat*) is thought to have arisen due to the influence of the following numeral, Common Slavonic \*desěti, ‘ten’ PIE \*dekṃ(t) (Comrie 1992:760). Winter (1969) discusses how the form for ‘four’ influences ‘five’ in Indo-European languages. These examples illustrate that ‘[a]nalogy is a powerful factor in counting, in both alliteration and rhyme, such that regular sound laws can be broken.’ (Sidwell 1999:256).

<sup>5</sup> Liquid-stop metathesis has occurred in Blagar-Bama *akur* ‘two’, but not in other Blagar dialects.

<sup>6</sup> In Abui and Kamang, the vowels in ‘four’ display some irregular patterns. For Kamang, it is necessary to posit the following metathesis: Proto-Alor Pantar \*buta < \*bita < biat.

<sup>7</sup> For instance, pAP \*bui ‘betel nut’ is probably borrowed from Austronesian (proto-West Malayo-Polynesian) \*buyuq ‘leaf of betel vine’ (Blust nd).

<sup>8</sup> The function of Malay as the *lingua franca* of the Dutch East Indies appears irrelevant for Alor-Pantar, as the area was under (remote) Portuguese control till 1860, and Dutch colonial influence only became apparent in the first decades of the 20<sup>th</sup> century (Klamer 2010:14 and references cited there). There is no evidence that Malay was used as a trade language in the Alor archipelago in pre-colonial times. On the other hand, there is anecdotal evidence that Alorese was used for interethnic communications in Pantar and coastal parts of west Alor until the mid 20<sup>th</sup> century (Klamer 2011).

Table 2: AP numerals ‘one’ to ‘five’

		‘one’	‘two’	‘three’	‘four’	‘five’
Proto-AP		*nuk	*araku	*(a)tiga	*buta	*jiwesin
Pantar	Western Pantar	<i>anuku</i>	<i>alaku</i>	<i>atiga</i>	<i>atu</i>	<i>jasin</i>
	Deing	<i>nuk</i>	<i>raq</i>	<i>atig</i>	<i>ut</i>	<i>asan</i>
	Sar	<i>nuk</i>	<i>raq</i>	<i>tig</i>	<i>ut</i>	<i>jawan</i>
	Teiwa	<i>nuk</i>	<i>haraq ~ raq</i>	<i>jerig</i>	<i>ut</i>	<i>jusan</i>
	Kaera	<i>nuk(u)</i>	<i>(a)rax-</i>	<i>(i/u)tug</i>	<i>ut</i>	<i>isim</i>
Straits	Blagar-Bama <sup>5</sup>	<i>nuku</i>	<i>akur</i>	<i>tuge</i>	<i>ut</i>	<i>isin</i>
	Blagar-Dolabang	<i>nu</i>	<i>aru</i>	<i>tue</i>	<i>buta</i>	<i>isin</i>
	Reta	<i>anu</i>	<i>alo</i>	<i>atoga</i>	<i>buta ~ wuta</i>	<i>avehan</i>
	Kabola	<i>nu</i>	<i>olo</i>	<i>towo</i>	<i>ut</i>	<i>iwesen</i>
W Alor	Adang	<i>nu</i>	<i>alo</i>	<i>tuo</i>	<i>ut</i>	<i>ifihin</i>
	Hamap	<i>nu</i>	<i>alo</i>	<i>tof</i>	<i>ut</i>	<i>ivehin</i>
	Klon	<i>nuk</i>	<i>orok</i>	<i>ton</i>	<i>ut</i>	<i>eweh</i>
	Kui	<i>nuku</i>	<i>oruku</i>	<i>siwa</i>	<i>usa</i>	<i>jesan</i>
C & E Alor	Abui	<i>nuku</i>	<i>ajoku</i>	<i>sua</i>	<i>buti</i>	<i>jetin</i>
	Kamang	<i>nok</i>	<i>ok</i>	<i>su</i>	<i>biat</i> <sup>6</sup>	<i>iwesen</i>
	Sawila	<i>(sundana)</i> <sup>†</sup>	<i>jaku</i>	<i>tuo</i>	<i>(ara:si:ku)</i>	<i>jo:tin</i>
	Kula	<i>(sona)</i>	<i>jakwu</i>	<i>tu</i>	<i>(arasiku)</i>	<i>jawatena</i>
	Wersing	<i>no</i>	<i>joku</i>	<i>tu</i>	<i>(arasoku)</i>	<i>wetin</i>

<sup>†</sup>Brackets indicate forms not reflecting proto-Alor-Pantar reconstructions.



pAP. If there were an Austronesian numeral ‘three’ borrowed into the family, this would more likely be similar to proto-Austronesian \**telu* ‘three’ (Blust 2009:268) instead of Malay *tiga*. The Austronesian languages surrounding Alor-Pantar reflect proto-Austronesian \**telu*. For instance, Alorese (an Austronesian language spoken on the coasts of Pantar and Alor) has *tilu*, Kedang (on Lembata) has *telu*, the language of Atauro (a small island of the north coast of Timor) has *hetelu* and Tetun Fehan (on Timor) has *tolu*.

In short, AP languages have by and large cognate forms for numerals ‘one’ to ‘five’ that reflect monomorphemic lexemes inherited from proto-AP.

## 5 Numerals ‘six’ to ‘nine’

Unlike numerals ‘one’ to ‘five’ which show significant stability across the AP group, numerals ‘six’ to ‘nine’ have undergone several changes in their history. In most AP languages, ‘six’ is morphologically simple, but in a subset of the languages bi-morphemic ‘six’ [5+1] has been innovated (§ 5.1). The numerals ‘seven’, ‘eight’, and ‘nine’ show more complex histories, with some being historically constructed by addition to a quinary base [5+*n*] and others by subtraction from a decimal base [10-*n*] (§ 5.2).

### 5.1 Numeral ‘six’: Simplex and compound forms

In most AP languages, the form ‘six’ is mono-morphemic, as shown in Table 3. Four languages have a compound ‘six’: Western Pantar in the west, and Alor, Sawila and Kula in the east. Kamang and Wersing display both patterns across their dialects: Kamang-Takailubui and Wersing-Pureman have simplex ‘six’, while Kamang-Atoitaa and Wersing-Kolana have compound ‘six’.

The simplex numeral ‘six’ reconstructs as pAP \**talam* ‘six’. It is generally assumed that a base-five system originates from counting the fingers of one hand. In such a system, the numeral ‘six’ often involves crossing over from one hand to the other,<sup>9</sup> and may etymologically be related to words like ‘cross over’ (Majewicz 1981, 1984, Lynch 2009: 399-401). Synchronic evidence that pAP \**talam* may have been such a ‘cross-over’ verb comes from Sawila, which has a modern form *talamanj* ‘step on, change legs in dance’.

---

<sup>9</sup> Cross-linguistically, other strategies to express ‘six’ include (in bodily counting routines) touching or grabbing the wrist (Evans 2009, Donohue 2008), or using the etymon ‘fist’ (Plank 2009: 343). We do not find any such practices in Alor-Pantar.

Table 3: AP numerals for ‘six’

		simplex ‘six’	compound ‘six’
Proto-AP		*talam	
Pantar	Western Pantar		<i>hisnakkun</i>
	Deing	<i>talan</i>	
	Sar	<i>tejan</i>	
	Teiwa	<i>tia:m</i>	
	Kaera	<i>tia:m</i>	
Straits	Blagar-Bama	<i>tajan</i>	
	Blagar-Dolabang	<i>talin</i>	
	Reta	<i>talaun</i>	
W Alor	Kabola	<i>talan</i>	
	Adang	<i>talan</i>	
	Hamap	<i>talan</i>	
	Klon	<i>tlan</i>	
	Kui	<i>talama</i>	
C & E Alor	Abui	<i>tala:ma</i>	
	Kamang-Takailubui	<i>ta:ma</i>	
	Kamang-Atoitaa		<i>isiɲnok</i>
	Sawila		<i>jo:tiɲsundana</i>
	Kula		<i>jawatensona</i>
	Wersing-Pureman	<i>təlam</i>	
	Wersing-Kolana		<i>wetiɲnun</i>

AP compounds for ‘six’ are composed of two juxtaposed numeral morphemes. In several AP languages, compounds for ‘six’ have replaced etymological \*talam ‘six’. There appears to have been three independent innovations of this kind. One area where this has happened is eastern Alor, represented by Sawila, Kula and Wersing in Table 4. Kula, Sawila and Wersing-Kolana have replaced \*talam ‘six’ with a base-five compound, as set out in (3). Whereas Sawila and Kula use ‘one’ in the compounds, the morpheme *nun* used in Wersing-Kolana for the ‘[plus] one’ part of the compound is not identical to the synchronic numeral ‘one’, which is *no*. Rather, *nun* appears to be a reflex of a distinct pAP lexeme \*nakun ‘single’, which is reflected in, for example, Kamang *nukun* ‘single’, and Western Pantar *nakkin* ‘single’.

(3) Formation of ‘six’ in eastern Alor languages

Sawila:

*jo:tiŋsundana* ‘six’ < *jo:tiŋ* ‘five’ [plus] *sundana* ‘one’

Kula:

*jawatensona* ‘six’ < *jawetena* ‘five’ [plus] *sona* ‘one’

Wersing-Kolana:

*wetiŋnun* ‘six’ < *wetiŋ* ‘five’ [plus] *nun* ‘single’

As these three languages are in close contact with each another, it is likely that the innovative use of a base-five compound for ‘six’ has diffused among them. This has probably happened relatively recently, since the members of the compounds are transparently related to existing cardinals. Older compounds show more divergence between the compound members and the individual numerals these derive from (cf. the formally less transparent base-five compounds discussed in § 5.2.1). The view that the transparent base-five forms are innovations is confirmed by the fact that Wersing-Pureman, the dialect spoken in –what according to oral traditions is– the Wersing homeland, preserves etymological, simplex ‘six’: *talam*.

In the north-central Alor language Kamang, the formation of ‘six’ differs between dialects, as set out in (4). The Atoitaa dialect has a base-five compound of ‘five [plus] one’, while the dialect of Atoitaa reflect pAP \*talam ‘six’.

(4) Formation of ‘six’ in Kamang dialects

Kamang-Atoitaa: *isiŋnok* ‘six’ < *iweŋ* ‘five’ [plus] *nok* ‘one’

Kamang-Takailubui: *ta:ma* ‘six’

Other language varieties in central Alor (Suboo, Tice, Moo and Manetaa), also have *ta:ma* ‘six’. The dominance of etymological ‘six’ in the area indicates that the

Atoitaa Kamang pattern is a recent innovation, probably occurring by extending the base-five pattern used in forming ‘seven’ through ‘nine’ to also include ‘six’.

In contrast to the transparent additive compounds found in the languages of eastern and north-central Alor, Western Pantar ‘six’ is structured more opaquely. There are two morphemes in Western Pantar *hisnakkun* ‘six’: (i) *his-*, a morpheme which has no independent meaning and; (ii) *-nakkun*, a morpheme originating in the Western Pantar verb *nakkin* ‘be single, alone’ (< pAP \*nakung ‘single’). The two morphemes are still apparent in the distributive form of ‘six’, *hisnakkun~nakkun* ‘six~REDUP’ ‘six by six’, where the second element reduplicates, contrasting with the distributive of monomorphemic numerals, e.g., *alaku~alaku* ‘two~REDUP’ ‘two by two’ (Klamer et al. this volume). The initial *his-* morpheme of the compound appears to be a borrowing of an Austronesian numeral ‘one’ (< pAN \*esa ~ isa). Initial [h] in the Western Pantar form *his-* is a non-phonemic consonant that appears before /i/, so that the underlying phonological form of *his-* is in fact /is-/. This matches well with the forms of the ‘one’ numeral in many nearby Austronesian languages on Flores (e.g. Nage *esa* ‘one’) and Timor (e.g. Tokodede *iso* ‘one’).

The distinct elements of the Western Pantar compound ‘six’ indicate that this numeral must have developed independently from the base-five forms for ‘six’ as found in central-east Alor languages. It appears that Western Pantar *hisnakkun* represents a partial calque of the base-five pattern found in Austronesian languages of Timor. In Tokodede and Mambae, ‘six’ is formed as ‘five-and-one’: Mambae *lim-nain-ide*, Tokodede *lim-woun-iso*. However, the initial *lim* ‘five’ is typically dropped, leaving simply ‘and-one’ to denote ‘six’: Mambae *nain-ide*, Tokodede *woun-iso*. Western Pantar *hisnakkun* may have borrowed Austronesian ‘one’ for the first half (*his*), while for the second half it uses a native element meaning ‘single’. The resulting combination ‘one-single’ is, then, a mediation of numeral constructions from different languages.

## 5.2 Numerals ‘seven’ to ‘nine’

The AP languages invariably have compound forms for ‘seven’, ‘eight’, ‘nine’, ‘ten’ and the decades. The compounds are constructed in two distinct ways. One is the additive base-five compound [5+n] (§ 5.2.1), the second a subtractive base-ten compound [10-n] (§ 5.2.2).

Table 4: Numerals ‘seven’ to ‘nine’ in Pantar and central-east Alor

		‘seven’ 5 2	‘eight’ 5 3	‘nine’ 5 4
Pantar	Deing	<i>jewasrak</i>	<i>santig</i>	<i>sanut</i>
	Sar	<i>jisraq</i>	<i>jinatig</i>	<i>jinaut</i>
	Teiwa	<i>jesraq</i>	<i>jesnerig</i>	<i>jesnaʔut</i>
	Kaera	<i>jesrax-</i>	<i>jentug</i>	<i>jeniut</i>
C&E Alor	Abui	<i>jetinjajoku</i>	<i>jetinjsua</i>	<i>jetinjbuti</i>
	Kamang-Takailubui	<i>wesinjok</i>	<i>wesinjsu</i>	<i>wesinjbiat</i>
	Kamang-Atoitaa	<i>isinjok</i>	<i>isinjsu</i>	<i>isinjbiat</i>
	Sawila	<i>jo:tinjaku</i>	<i>jo:tinjtuo</i>	<i>jo:tinjara:siiku</i>
	Kula	<i>jawatenjakwu</i>	<i>jawatentu</i>	<i>jawatenarasiku</i>
	Wersing	<i>wetinjoku</i>	<i>wetintu</i>	<i>wetinarasoku</i>

### 5.2.1 Additive base-five compounds

Numerals ‘seven’ to ‘nine’ that are historically formed as additive base-five numeral (i.e., [5 2] ‘seven’, [5 3] ‘eight’, [5 4] ‘nine’) are found in both Pantar and central-east Alor. Table 4 presents an overview.

The languages of central-east Alor construct ‘seven’ to ‘nine’ with an additive base-five system that is synchronically transparent. That is, speakers of these languages readily parse their numerals ‘seven’ to ‘nine’ as being composed of the synchronic numeral ‘five’ followed by ‘two’, ‘three’, or ‘four’, as set out in (5).

#### (5) Compound ‘seven’ to nine’ in central-east Alor

a.	‘seven’	< ‘five’	[plus] ‘two’
Abui	<i>jetinjajoku</i>	< <i>jetinj</i>	<i>ajoku</i>
Kamang (T)	<i>wesinjok</i>	< <i>wesinj</i>	<i>ok</i>
Kamang (A)	<i>isinjok</i>	< <i>iwesinj</i>	<i>ok</i>
Sawila	<i>jo:tinjaku</i>	< <i>jo:tinj</i>	<i>jaku</i>
Kula	<i>jawatenjakwu</i>	< <i>jawatena</i>	<i>jakwu</i>
Wersing	<i>wetinjoku</i>	< <i>wetinj</i>	<i>joku</i>

b.	‘eight’	< ‘five’	[plus] ‘three’
Abui	<i>jetɪŋsua</i>	< <i>jetɪŋ</i>	<i>sua</i>
Kamang (T)	<i>wesɪŋsu</i>	< <i>wesɪŋ</i>	<i>su</i>
Kamang (A)	<i>isɪŋsu</i>	< <i>iwesɪŋ</i>	<i>su</i>
Sawila	<i>jo:tiŋtuo</i>	< <i>jo:tiŋ</i>	<i>tuo</i>
Kula	<i>jawatentu</i>	< <i>jawatena</i>	<i>tu</i>
Wersing	<i>wetiŋtu</i>	< <i>wetiŋ</i>	<i>tu</i>
c.	‘nine’	< ‘five’	[plus] ‘four’
Abui	<i>jetɪŋbuti</i>	< <i>jetɪŋ</i>	<i>buti</i>
Kamang (A)	<i>isɪŋbiat</i>	< <i>iwesɪŋ</i>	<i>biat</i>
Kamang (T)	<i>wesɪŋbiat</i>	< <i>wesɪŋ</i>	<i>biat</i>
Sawila	<i>jo:tiŋara:si:ku</i>	< <i>jo:tiŋ</i>	<i>ara:si:ku</i>
Kula	<i>jawatenarasiku</i>	< <i>jawatena</i>	<i>arasiku</i>
Wersing	<i>wetiŋarasoku</i>	< <i>wetiŋ</i>	<i>arasoku</i>

By contrast, languages of the Pantar subgroup have compounds for numerals ‘seven’ through ‘nine’ that are synchronically non-transparent. The patterns found across the Pantar languages show certain regularities, indicating that the reductions were probably already present in their immediate ancestor, proto-Pantar (pP).<sup>10</sup> In (6) we present the reconstructed numeral compounds and their constituent numeral roots. Only the major changes leading to the forms in the modern languages are detailed here.

The final segments /in/ of pP \*jiwasin ‘five’ were already lost in pP ‘seven’. This was followed by the loss of (probably unstressed) medial /wa/ in the subgroup containing Sar, Teiwa and Kaera (proto-Central Pantar, pCP). In the pP-forms of ‘eight’ and ‘nine’, medial /we/ of \*jewasin ‘five’ was lost, as well as the final /a/ of \*(a)tiga ‘three’. In pCP ‘eight’ and ‘nine’, a metathesis of /an/ to /na/ occurred,<sup>11</sup> followed by a reduction of the resulting /sn/ cluster to /s/ in the group containing Sar and Kaera (proto-Central East Pantar, pCEP).

<sup>10</sup> The sub-groups within the AP group that we name here are based purely on evidence from formal and phonological (often sporadic and/or irregular) changes shared between languages in their numerals. The reconstruction of pAP in Holton et al. (2012) is too preliminary and coarse-grained to pick up any real subgrouping evidence. We take the detailed study of numerals that we make here to be indicative of how we may go about identifying AP subgroups in the future.

<sup>11</sup> The medial glottal stop in the Teiwa form appears to have been inserted as a syllable boundary marker as the adjacent vowels /au/ of pCP \*jesnaut began to harmonise in Teiwa.

- (6) Developments in proto-Pantar cardinals ‘seven’ to ‘nine’
- a. ‘seven’:  
pP \*jewasin ‘five’ [plus] \*raqo ‘two’ > \***ewas**raqo > pCP \*jesraqo
  - b. ‘eight’:  
\*jewasin ‘five’ [plus] \*atiga ‘three’ > \*je’santig > pCP \*jesnatig > pCEP  
\*jenatig
  - c. ‘nine’:  
pP \*jewasin ‘five’ [plus] \*ut ‘four’ > \*je’sanut > pCP \*jesnaut > pCEP  
\*jenaut

The presence of base-five numerals for ‘seven’ through ‘nine’ in separate groups of the AP languages at opposite ends of the archipelago, coupled with the absence of any other equally widely attested forms for these numerals, is a strong indication that a base-five system was used in proto-AP ‘seven’ through ‘nine’.

The difference between base-five numerals in Pantar and central-east Alor languages is merely in the transparency of formatives in the compound numerals. While in the Pantar group base-five compounds have been reduced to such an extent that they are no longer transparent, the central-east Alor languages have retained base-five as a transparent and productive system.

### 5.2.2 Subtractive base-ten compounds and extensions

The second strategy of creating numerals ‘seven’ through ‘nine’ found in the AP languages is subtraction, that is, [10-3] for ‘seven’, [10-2] for ‘eight’ and [10-1] for ‘nine’. This strategy is found in the Straits and West Alor languages. Table 5 sets out the numerals under discussion in this group along with their reconstructed forms in the ancestor language proto-Straits-West-Alor (pSWA). The final two language in Table 5, Kui and Western Pantar, have innovated their own distinctive forms as indicated by the brackets. They are nevertheless included here because, as will be seen in § 5.2.3, they both include some of the formatives distinctive of pSWA numerals ‘seven’ to ‘nine’.

The subtractive basis for the formation of Straits-West-Alor numerals ‘seven’ through ‘nine’ is evident from the remnants of reflexes of proto-AP \*(a)tiga ‘three’, \*araqo ‘two’ and \*nuk ‘one’ in the final syllables of modern forms, as set out in (7). Bolding indicates the matching strings of segments.

Table 5: Numerals ‘seven’ to ‘nine’ in the Straits and West Alor

		‘seven’	‘eight’	‘nine’
proto-Straits-West-Alor		* <b>futitoga</b>	* <b>turarok</b>	* <b>tuka'rinuk</b>
		7 3	[10]-2	[10]-1
Straits	Blagar-Bama	<i>titu</i>	<i>tuakur</i>	<i>tukurunuku</i>
	Blagar-Dolabang	<i>fititu</i>	<i>tuaru</i>	<i>turinu</i>
	Reta	<i>bititoga</i>	<i>tulalo</i>	<i>tukanu</i>
West Alor	Kabola	<i>wutito</i>	<i>turlo</i>	<i>ti?inu</i>
	Adang	<i>ititɔ</i>	<i>turlo</i>	<i>ti?enu</i>
	Hamap	<i>itito</i>	<i>turalo</i>	<i>tieu</i>
	Klon	<i>uson</i>	<i>tidorok</i>	<i>tukainuk</i>
South-West Alor	Kui	( <i>jesaroku</i> )	( <i>tadusa</i> )	( <i>jesanusa</i> )
Pantar	Western Pantar	( <i>betalaku</i> )	( <i>betiga</i> )	( <i>anukutannan</i> )

(7) Formatives ‘three’, ‘two’, and ‘one’ in Straits-West-Alor ‘seven’ through ‘nine’

a. ‘seven’ Compare: ‘three’

Blagar-B:	<i>titu</i>	<i>tuge</i>
Blagar-D:	<i>fititu</i>	<i>tue</i>
Reta:	<i>bititoga</i>	<i>atoga</i>
Kabola:	<i>wutito</i>	<i>towo</i>
Adang:	<i>ititɔ</i>	<i>tuo</i>
Hamap:	<i>itito</i>	<i>tof</i>
Klon:	<i>uson</i>	<i>ton</i>

b. ‘eight’ Compare: ‘two’

Blagar-B:	<i>tuakur</i>	<i>akur</i>
Blagar-D:	<i>tuaru</i>	<i>aru</i>
Reta:	<i>tulalo</i>	<i>alo</i>
Kabola:	<i>turlo</i>	<i>alo</i>
Adang:	<i>turlo</i>	<i>alo</i>
Hamap:	<i>turalo</i>	<i>alo</i>
Klon:	<i>tidorok</i>	<i>orok</i>



c.	‘nine’	Compare: ‘one’
Blagar-B:	<i>tukurunuku</i>	<i>nuku</i>
Blagar-D:	<i>turinu</i>	<i>nu</i>
Reta:	<i>tukanu</i>	<i>anu</i>
Kabola:	<i>ti?inu</i>	<i>nu</i>
Adang:	<i>ti?enu</i>	<i>nu</i>
Hamap:	<i>tieu</i>	<i>nu</i>
Klon:	<i>tukainuk</i>	<i>nuk</i>

Two different subtractive bases are apparent in the modern forms: (i) in ‘eight’ and ‘nine’, there is a synchronically unanalysable initial morpheme, which is followed by reflexes of ‘two’ and ‘one’, and; (ii) in ‘seven’, we see an augend that we argue below to be a borrowed reflex of proto-Austronesian \*pitu ‘seven’, with reflexes of ‘three’ as addend. We discuss these two constructions now in turn.

The unanalysable initial elements in the compounds for ‘eight’ (\*tur-) and ‘nine’ appear to go back to a single morpheme pre-proto-Straits-West-Alor \*tukari, originally meaning something like ‘[ten] less’ or ‘[ten] take away’. On this reconstruction, pre-pSWA \*tukari was already reduced to \*tur- in pSWA ‘eight’, but maintained in ‘nine’. We suggest that \*tukari meant ‘less’ or ‘take away’ rather than ‘ten’ for two reasons. First, the reconstructed pAP \*qar ‘ten’ (Holton et al. 2012) has a distinct form which cannot be reconciled with pre-pSWA \*tukari. Second, to assign \*tukari the meaning ‘ten’ would imply that the numerals formed by subtraction would be composed of a simple sequence of the subtrahend and the minuend. This would be a cross-linguistically unusual pattern and is judged to be unlikely here, but by no means impossible.

We analyse these subtractive numerals as originally constructed along the lines of ‘ten less one’, ‘ten less two’, and ‘ten less three’. However, over time, the numeral overtly denoting ‘ten’ was dropped and ‘less one’ was conventionalised to mean ‘nine’, ‘less two’ to mean ‘eight’, and ‘less three’ to mean ‘seven’. In turn, it appears that pre-pSWA \*tukari was reanalysed as a subtrahend rather than the actual morpheme expressing the subtraction. This is seen in its replacement by another base in pSWA ‘seven’, the other subtrahend that is apparent in the modern numerals, \*buti-. We propose that this is a borrowed numeral which is a reflex of proto-Austronesian \*pitu ‘seven’. It is followed by reflexes of pAP \*(a)tiga ‘three’ to denote ‘seven’, and has replaced the pre-pSWA morpheme \*tukari in the numeral ‘seven’, as laid out in Table 6.

In other words, Proto-Straits-West-Alor \*butitoga is composed of \*buti, a borrowed base that is a reflex of PAN \*pitu ‘seven’, conjoined with *toga* as a reflex of pAP \*(a)tiga ‘three’. The Straits-West-Alor languages are located along a nar-

Table 6: Pre-Proto-Straits-West-Alor and Proto-Straits-West-Alor ‘seven’ to ‘nine’

		‘seven’	‘eight’	‘nine’
Pre-pSWA	stage I	<i>*tukaritoga</i> less.three	<i>*tukariarok</i> less.two	<i>*tukarinuk</i> less.one
	stage II	<i>*butitoga</i> seven.three		
pSWA		<i>*butitoga</i>	<i>*turarok</i>	<i>*tukarinuk</i>

row and busy strait where language contact with Austronesian speakers is highly plausible. The motivation for borrowing an Austronesian base for ‘seven’ may have been Austronesian cultural influence. Among Austronesian groups in eastern Indonesia, ‘seven’ is a culturally significant numeral (e.g., in Flores (Forth 2004:221); Kedang on Lembata (Barnes 1982:14-18); Tetun Fehan on Timor (Van Klinken 1999:102) and Kambara on east Sumba (Forth 1981:212-213)).

The resulting proto-Straits-West-Alor numeral compound ‘seven’ was, however, a mediation of the contact and the native numeral. By borrowing the numeral for ‘seven’, the Austronesian pattern was emulated, but by maintaining the original minuend ‘three’ along with the new Austronesian numeral functioning as subtrahend, the native Straits-West-Alor subtractive pattern was also partially preserved.

Such a rearrangement in which a numeral is formed mathematically incorrectly may appear unusual, but parallels are found in other languages of the area. For instance, in the Manufahi dialect of Bunaq (a language related to the Alor-Pantar languages spoken on Timor (Schapper 2010)), ‘six’ is denoted by *tomol-uen*, a compound of etymological ‘six’ and ‘one’. Bunaq-Manufahi is spoken in an area dominated by speakers of the Austronesian language Mambae and all Bunaq-Manufahi speakers also speak Mambae. As discussed in § 5.1, Mambae has a quinary system for the formation of numerals ‘six’ to ‘nine’. The Bunaq-Manufahi pattern of forming ‘six’ mediates between the native Bunaq pattern of *tomol* ‘six’ and Mambae *lim-nain-ide* ‘five-and-one’ > ‘six’, by combining *tomol* ‘six’ and *uen* ‘one’.

In an alternative analysis, proto-Straits-West-Alor *\*butitoga* ‘seven’ would be a compound of proto-Alor-Pantar *\*buta* ‘four’ and *\*(a)tiga* ‘three’ [4 3]. The advantage of this etymology is that no borrowing from Austronesian is invoked.

However, the analysis implies that proto-Straits-West-Alor innovated a numeral with a quaternary base as the initial member of an additive compound. This would effectively add a completely new (fourth) type to the structural inventory of numeral system types found in AP numerals.

Recall that AP numerals have (i) additive compounds with quinary bases as first element (e.g. 5[+2]), (ii) multiplicative compounds with quaternary bases as second (*not* first) element [e.g. 2[x]4], and (iii) subtractive compounds ([10]–2). While we cannot exclude the possibility that a (proto-)language invents a completely new structural type for a single numeral, we believe this scenario to be less likely than the borrowing plus reanalysis scenario outlined above.

It should be added that there is no evidence that a new [4 3] pattern could have been borrowed from neighbouring Austronesian language(s), as [4 3] ‘seven’ is not attested anywhere in the region (Schapper & Hammarström 2013). Numerals with a quaternary base are found in the region, but these are all multiplicative forms: [2 4] (Flores) or [4 2] (Lembata) ‘eight’ (see Table 13 and Appendix A.2).

### 5.2.3 Other mixed systems for ‘seven’ to ‘nine’

In the previous section it was mentioned that ‘seven’ through ‘nine’ in Kui and Western Pantar include formative elements of the Straits West Alor system. However, the formatives are part of different systems, using a range of bases.

In (8) we set out the formatives found in Kui ‘seven’ to ‘nine’. We can see that Kui has replaced the proto-Straits-West-Alor subtractive numerals for ‘seven’ and ‘nine’ with additive base-five numerals [5 2], [5 4]. The numeral *tadusa* ‘eight’ follows a different pattern, apparently being built from two morphemes: (i) the first element *tad-* appears to reflect the subtractive morpheme \*tur- (< pre-pSWA \*tukari) used in forming pSWA \*turarok ‘eight’ (see Table 6), and (ii) the second element *usa* is the Kui numeral ‘four’ (< pAP \*buta ‘four’).

#### (8) Formatives in Kui ‘seven’ to ‘nine’

- a. ‘seven’:  
*jesaroku* < *jesan* ‘five’ [plus] *oruku* ‘two’
- b. ‘eight’:  
*tadusa* < *tad-* *usa* ‘four’
- c. ‘nine’:  
*jesanusa* < *jesan* ‘five’ [plus] *usa* ‘four’

It appears that in Kui ‘eight’ has been imperfectly remodelled on a multiplication pattern ‘two [times] four’ [2x4]. The original proto-Straits-West-Alor \*tu-

rarak ‘eight’, historically composed of pre-PSWA \*tukari ‘less’ and pre-PSWA \*arok ‘two’, appears to have been reduced and reanalysed from subtractive ‘minus two’ to multiplicative ‘[two] times’. This new base was then combined with *usa* ‘four’ to reach ‘eight’. The /d/ in Kui *tadusa* ‘eight’ appears to have arisen through liquid dissimilation of the two /r/’s in the adjacent syllables. That is, as we see also in Klon *tidorok* ‘eight’, dissimilation applied such that \*turarak took on a hypothetical form like \*\*tudarok. In the history of Kui, the \*arok element of hypothetical \*\*tudarok was then replaced with *usa* ‘four’ to create \*\*tudusa ‘eight’ with the vowel changes u > o > a leading to modern Kui *tadusa* ‘eight’.

In (9) we set out the formatives found in Western Pantar ‘seven’ to ‘nine’. We can see that proto-Straits-West-Alor subtractive numeral \*tukarinuk ‘nine’ have been replaced by an innovative, but still subtractive form composed of the numeral ‘one’ denoting the subtrahend and the lexical verb ‘take away’ signalling the subtraction. The numerals ‘seven’ and ‘eight’ follow a different, innovative pattern in which *be-* ~ *bet-*, reflecting \*buti- as also used in the formation of proto-Straits-West-Alor \*butitoga ‘seven’, is combined with ‘two’ and ‘three’ to form ‘eight’ and ‘nine’ respectively.

(9) Formatives in Western Pantar ‘seven’ to ‘nine’

- a. ‘seven’:  
*betalaku* < *bet-* ‘?’ *alaku* ‘two’
- b. ‘eight’:  
*betiga* < *be-* ‘?’ *tiga* ‘three’
- c. ‘nine’:  
*anukutannan* < *anuku* ‘one’ *tannan* ‘take away’

Synchronic evidence for their poly-morphemic status includes the distributive formation of ‘seven’, which takes the right-most element as the base for the reduplication Klammer et al. (this volume): *betalaku*~*talaku* ‘seven~REDUP’ ‘seven by seven’, *betiga*~*tiga* ‘eight~REDUP’ ‘eight by eight’, and *anuktannan*~*tannan* ‘nine~REDUP’ ‘nine by nine’. The segmentation in distributives appears to be a historical relic. This is suggested by the irregularities in the distributive derivation of ‘seven’: there has been a reanalysis of the morpheme boundary between *bet-* and *alaku* ‘two’ to become *be-talaku*, analogous to the segmentation of the numeral *betiga* ‘three’. The reanalysis points to speakers not being able to decompose the complex numerals into their original forms.

Thus, in Western Pantar, \*buti- has been adopted not as a minuend (as in proto-Straits-West-Alor ‘seven’), but as an augend. We posit that proto-Western Pantar

Table 7: Proto-W Pantar developments leading to modern W Pantar ‘seven’, ‘eight’

		‘seven’	‘eight’
Proto-WP	stage I	<b>*jasinalaku</b> <b>five.two</b>	<b>*jasinatiga</b> <b>five.three</b>
	stage II	<b>*butialaku</b> <b>seven.two</b>	
	stage III		<b>*butiatoga</b> <b>seven.three</b>

originally had an additive base-5 system in which ‘seven’ and ‘eight’ were formed by means of compounds of ‘five [plus] two’ and ‘five [plus] three’ respectively, as set out in stage I in Table 7. In stage II, pre-Western Pantar ‘five’ is replaced by a reflex of proto-Austronesian \*pitu ‘seven’ borrowed either directly from an Austronesian language under the same forces for pre-proto-Straits-West-Alor as described in § 5.2.2, or perhaps more likely from a ([pre]-proto)-Straits-West-Alor language. In stage III, the pattern of using \*buti- as an augend for the formation of ‘seven’ in stage II is extended to the formation of ‘eight’ (Table 7).

In sum, in Kui, the proto-Straits-West-Alor subtractive numerals for ‘seven’ and ‘nine’ have been replaced by additive base-five forms [5 2] and [5 4], while ‘eight’ has become a base-four compound [2x4]. In Western Pantar, on the other hand, the proto-Straits-West-Alor subtractive numeral ‘nine’ has been replaced by an innovative, but still subtractive form composed of the numeral ‘one’ denoting the subtrahend and the lexical verb ‘take away’. Western Pantar ‘seven’ and ‘eight’ involve a borrowed and reanalysed quinary base.

## 6 Numerals ‘ten’ and above

### 6.1 Numeral ‘ten’: multiplied base-ten compound

Table 8 presents the numerals 10, 20 and 30. A decimal base \*qar ‘ten’ is reconstructable to proto-Alor-Pantar. This is reflected across AP languages, with the exception of central-eastern Alor, where languages eastwards of Kamang reflect innovative proto-Central-East-Alor (pCEA) \*adajaku ‘ten’. This form indicates that a quinary base may have at some point replaced the decimal base in these languages: the second element of the compound \*adajaku ‘ten’ is homophonous

Table 8: Numerals ‘ten’ and the formation of decades

		‘ten’	‘twenty’	‘thirty’
Pantar	Western Pantar	<i>ke†anuku</i>	<i>ke alaku</i>	<i>ke atiga</i>
	Deing	<i>qar nuk</i>	<i>qar raq</i>	<i>qar atig</i>
	Sar	<i>qar nuk</i>	<i>qar raq</i>	<i>qar tig</i>
	Teiwa	<i>qa:r nuk</i>	<i>qa:r raq</i>	<i>qa:r jerig</i>
	Kaera	<i>xar nuko</i>	<i>xar raxo</i>	<i>xar tug</i>
Straits	BlagarBama	<i>qar nuku</i>	<i>qar akur</i>	<i>qar tuge</i>
	BlagarDolabang	<i>?ari nu</i>	<i>?ari aru</i>	<i>?ari tue</i>
	Reta	<i>kara nu</i>	<i>kara alo</i>	<i>kara atoga</i>
West Alor	Kabola	<i>kar nu</i>	<i>kar ho(?)olo</i>	<i>kar towo</i>
	Adang	<i>?er nu</i>	<i>?er alo</i>	<i>?er tuo</i>
	Hamap	<i>air nu</i>	<i>air alo</i>	<i>air tof</i>
	Klon	<i>kar nuk</i>	<i>kar orok</i>	<i>kar toŋ</i>
	Kui	<i>kar nuku</i>	<i>kar oruku</i>	<i>kar siwa</i>
C&E Alor	Abui	<i>kar nuku</i>	<i>kar ajoku</i>	<i>kar sua</i>
	Kamang	<i>ata:k nok</i>	<i>ata:k ok</i>	<i>ata:k su</i>
	Sawila	<i>ada:ku</i>	<i>ada:ku maraku</i>	<i>ada:ku matua</i>
	Kula	<i>adajakwu</i>	<i>mijakwu</i>	<i>mitua</i>
	Wersing	<i>adajoku</i>	<i>adajoku mijoku</i>	<i>adajoku mitu</i>

†In Western Pantar the final consonant of \*qar underwent irregular loss at the word-internal morpheme boundary.

with \*jaku ‘2’ so that it appears to be composed as [(5?) x 2].

In the modern AP languages, reflexes of \*qar for the most part do not stand alone but must be combined with another numeral in order to signify. Decades (numerals denoting a set or series of ten such as ‘10’, ‘20’, ‘30’ etc.) in AP languages are typically formed by combining the decimal base with a multiplicand indicating the decade. Thus, ‘ten’ is composed of a reflex of \*qar and ‘one’ [10 1], ‘twenty’ of ‘ten’ and ‘two’ [10 2], ‘thirty’ of ‘ten’ and ‘three’ [10 3], and so on for higher decades.

In the east of Alor we find deviations from this majority pattern for forming decades. First, Sawila, Kula and Wersing do not denote ‘ten’ by ‘ten [times] one’ [10 1] as elsewhere, but employ the numeral ‘ten’ alone without ‘one’. Second, in the formation of decades higher than ‘ten’, these languages do not simply juxtapose numerals to express multiplication of the base-ten, but mark it with

a prefix (Kula/Wersing *mi-* and Sawila *m(a)-*) on the multiplicand. These are verbal prefixes which have developed from the proto-AP postposition \**mi* ‘be in’ (Holton & Robinson this volume). Attached to numerals *mi-* ‘TIME’ derives frequency verbs such as ‘to do twice’ in the Alor languages Kamang and Klon.<sup>12</sup>

- (10) Kamang (Schapper, field notes)

*Alma uh ok an-iŋ=daŋ kai mi-ok.*  
 person CLF two thus-SET=WHEN cheer TIME-two  
 ‘Two person(’s heads) means (we) cheer twice.’

- (11) Klon (Baird 2008)

*... mid beh go-duur o mi-orok ...*  
 ... climb branch 3-cut DEM TIME-two ...  
 ‘... (he) climbed up (and) cut the branch twice...’

In the east of Alor, prefix *mi-* occurs on multiplicands to denote decades ‘twenty’ and above. This appears an extension of how *mi-* derives frequency verbs and ordinals from cardinals. In other words, the construction used to express higher decades in eastern Alor languages can be paraphrased as ‘ten twice’ for ‘twenty’, ‘ten thrice’ for ‘thirty’, and so on. In Kula, the use of *mi-* in the decade construction has conventionalised to such an extent that *adajakwu* ‘ten’ can be left off entirely, with the prefixed multiplicand carrying the decade meaning alone. That is, *mi-jakwu* for instance, would etymologically denote ‘twice’ or ‘second’, but is now used alone to mean ‘twenty’.

## 6.2 Numerals within decades

A ‘decade’ is a numeral which is a set or series of ten (e.g., ‘20’, ‘30’ etc.). The term ‘numeral within decades’ is used here to refer to any numeral expression such as ‘eighteen’, ‘eighty-nine’ etc., involving an operator word that signifies addition.<sup>13</sup> In AP languages, an additive operator separates the decades from the numerals ‘one’ to ‘nine’, as illustrated in Table 9.

The additive operator is not used to combine decades, hundreds or thousands with each other, as illustrated with ‘1999’ in several languages in (12).

<sup>12</sup> By contrast, *mi* derives ordinals in the Pantar-Straits languages Kaera, Blagar and Adang (Klamer et al. this volume).

<sup>13</sup> Such operators are referred to variously in the typological literature as ‘marker of addition’, ‘additive marker’ or ‘additive link’. See, e.g., Greenberg (1978:264-265); Hanke (2010:73).

Table 9: AP language compounds for ‘eighteen’

		‘ten’	‘one’	OPERATOR	‘eight’
Pantar	West Pantar	<i>ke</i>	<i>anuku</i>	<i>wali</i>	<i>betiga</i>
	Teiwa	<i>qa:r</i>	<i>nuk</i>	<i>rug</i>	<i>jesnerig</i>
	Kaera	<i>xar</i>	<i>nuk</i>	<i>beti</i>	<i>jentug</i>
Straits	Blagar-Bama	<i>qar</i>	<i>nuku</i>	<i>wali</i>	<i>tuakur</i>
	Blagar-Dolabang	<i>?ari</i>	<i>nu</i>	<i>belta</i>	<i>tuaru</i>
W Alor	Adang	<i>er</i>	<i>nu</i>	<i>faliŋ</i>	<i>turlo</i>
	Klon	<i>kar</i>	<i>nuk</i>	<i>awa</i>	<i>tidorok</i>
C&E Alor	Abui	<i>kar</i>	<i>nuku</i>	<i>wal</i>	<i>jetinjsua</i>
	Kamang	<i>ata:k</i>	<i>nok</i>	<i>wa:l</i>	<i>isinsu</i>
	Sawila	<i>ada:ku</i>		<i>garisiŋ</i>	<i>jo:tiŋtua</i>
	Kula	<i>adajakwu</i>		<i>arasiŋ</i>	<i>jawatentu</i>
	Wersing	<i>adajoku</i>		<i>weresiŋ</i>	<i>wetiŋtu</i>

- (12)
- |           |             |             |               |                  |            |                  |                 |                  |                 |   |
|-----------|-------------|-------------|---------------|------------------|------------|------------------|-----------------|------------------|-----------------|---|
|           | 1000        | 1           | 100           | 9                |            | 10               | 9               |                  | OPER            | 9 |
| Teiwa:    | <i>ribu</i> | <i>nuk</i>  | <i>ratu</i>   | <i>jesna?ut</i>  |            | <i>qa:r</i>      | <i>jesna?ut</i> | <i>rug</i>       | <i>jesna?ut</i> |   |
| W Pantar: | <i>ribu</i> |             | <i>ratu</i>   | <i>nuktannaŋ</i> | <i>ke</i>  | <i>nuktannaŋ</i> | <i>wali</i>     | <i>nuktannaŋ</i> |                 |   |
| Abui:     | <i>rifi</i> | <i>nuku</i> | <i>aisaha</i> | <i>jetinbuti</i> | <i>kar</i> | <i>jetinbuti</i> | <i>wal</i>      | <i>jetinbuti</i> |                 |   |
| Kamang:   | <i>ribu</i> | <i>nuk</i>  | <i>asaka</i>  | <i>isiŋbiat</i>  |            | <i>ata:k</i>     | <i>isiŋbiat</i> | <i>wa:l</i>      | <i>isiŋbiat</i> |   |

An additive operator \*wali(ŋ) can be reconstructed to proto-Alor-Pantar. In modern AP languages, the operator is for the most part a semantically empty lexeme without meaning outside of the numeral formula. However, some modern languages have homophonous lexical verb roots with semantics plausibly related to the additive operator: Teiwa and Kaera *wal* are verbs meaning ‘fill, full’, and Abui *wal-* is a verb meaning ‘gather more’. These might suggest that the proto-AP additive linker \*wali(ŋ) was a lexeme meaning ‘add, (do) again’.

Not all modern AP languages reflect the reconstructed operator. Kaera and Blagar-Dolabang have apparently related operators, while Teiwa has a unique form *rug*. In eastern Alor languages, the additive operators (Kula *arisiŋ*, Sawila *garisiŋ*, Wersing *weresiŋ*) are the result of shared borrowing from an Austronesian language of Timor, most likely Tokodede. The Austronesian languages of eastern Timor invariably use ‘more’ with a variant of the form /geresiŋ/ as the additive operator in numerals. Examples are provided in (13).



- (13) Additive operators in ‘eleven’ in the Austronesian languages of eastern Timor

	ten	OPERATOR	one
Tokodede	<i>sagulu geresi</i>		<i>iso</i>
Kemak	<i>sapulu resi</i>		<i>sia</i>
Tetun	<i>sanulu resin</i>		<i>ida</i>
Mambae	<i>sagul resi</i>		<i>kid</i>
Atauro	<i>senjulu resi</i>		<i>hea</i>

### 6.3 Multiples of ‘hundred’ and ‘thousand’

In most AP languages, bases for ‘hundred’ and ‘thousand’ cannot be used as numerals on their own. That is, they must be juxtaposed with a following multiplicand, so that ‘one hundred’ is [100x1], ‘two hundred’ [100x2] ‘200’, and so on (Table 10). A handful of languages do not conform to this pattern. Western Pantar *ratu*, Kui *asaga*, Kula *gasaka* and Wersing *aska* ‘hundred’ can be used independently to denote ‘100’. Western Pantar *ribu* is also able to independently denote ‘1000’, but may also appear with the unrelated form *je* to make *ribu je* ‘1000’. *Je* is also used in the ordinal ‘first’ (Klamer et al. this volume). Sawila *dana* and Kula *dena* are reductions of a different proto-form \*sundana ‘one’ (compare the forms for ‘one’ in Table 2).

Across much of Alor we find reflexes of a form \*a(j)saka ‘hundred’, but it is not clear to what level this form should be reconstructed. The languages of Pantar, Straits and West Alor have borrowed an Austronesian form reflecting PAN \*Ratus ‘hundred’.<sup>14</sup> There is no evidence of an indigenous AP numeral for ‘thousand’; all AP languages have borrowings from Austronesian reflecting PAN \*libu ‘thousands’. Possible source languages include Malay (*ratus* ‘hundreds’, *ribu* ‘thousands’), or Lamaholot (spoken on Adonara, Lomblen, Solor and Flores located west of Pantar), which employs the bases *ratu* ‘hundreds’ and *ribu* ‘thousands’. Austronesian languages in eastern Timor are not probable sources as they reflect neither PAN \*R in ‘hundred’ (e.g., Tetun, Kemak and Tokodede *atus* ‘hundred’) nor PAN \*b in ‘thousand’ (e.g., Tetun and Tokodede *rihun*, Kemak *lihur* ‘thousand’).

<sup>14</sup> In proto-Austronesian \*R represents an alveolar or uvular trill, contrasting with Proto-Austronesian \*r which is thought to have been an alveolar flap.

Table 10: Numerals with bases ‘100’ and ‘1000’

		‘100’	‘200’	‘1000’	‘2000’
Pantar	West Pantar	<i>ratu</i>	<i>ratu alaku</i>	<i>(a)ribu (ye)</i>	<i>ribu (alaku)</i>
	Deing	<i>aratu nuk</i>	<i>aratu raq</i>	<i>aribu nuk</i>	<i>aribu raq</i>
	Sar	<i>ratu nuk</i>	<i>ratu raq</i>	<i>ribu nuk</i>	<i>ribu raq</i>
	Teiwa	<i>ratu nuk</i>	<i>ratu (ha)raq</i>	<i>ribu nuk</i>	<i>ribu (ha)raq</i>
	Kaera	<i>ratu nuk</i>	<i>ratu rax-</i>	<i>ribu nuk</i>	<i>ribu rax-</i>
Straits	Blagar-Bama	<i>ratu nuk</i>	<i>ratu akur</i>	<i>ribu nuk</i>	<i>ribu akur</i>
	Blagar-Dolabang	<i>ratu nu</i>	<i>ratu aru</i>	<i>ribu nu</i>	<i>ribu aru</i>
	Reta	<i>ratu anu</i>	<i>ratu alo</i>	<i>ribu ano</i>	<i>ribu alo</i>
W Alor	Kabola	<i>rat nu</i>	<i>rat ho(?)olo</i>	<i>rib nu</i>	<i>rib ho(?)olo</i>
	Adang	<i>rat nu</i>	<i>rat alo</i>	<i>rib nu</i>	<i>rib alo</i>
	Hamap	<i>rat nu</i>	<i>rat alo</i>	— <sup>†</sup>	—
	Klon	<i>eska nok</i>	<i>eska orok</i>	—	—
C & E Alor	Kui	<i>asaga</i>	<i>asaga oruku</i>	<i>rab nuk</i>	<i>rab oruku</i>
	Abui	<i>aisaha nu</i>	<i>aisaha ajoku</i>	<i>rifi nuk</i>	<i>rifi ajoku</i>
	Kamang (L/U)	<i>asaka nok</i>	<i>asaka ok</i>	<i>libu nok</i>	<i>libu ok</i>
	Sawila	<i>asaka dana</i>	<i>asaka jaku</i>	<i>ri:bu dana</i>	<i>ri:bu jaku</i>
	Kula	<i>gasaka</i>	<i>gasaka jakwana</i>	<i>rib dena</i>	—
	Wersing	<i>aska</i>	<i>aska joku</i>	<i>ribu no</i>	—

† ‘—’ denotes that no data is available for this numeral

## 7 Alor-Pantar numerals from a typological and areal perspective

In this section, we consider how forms and systems used in the composition of AP numerals relate to those used in other languages. First, we place the AP numeral systems in a broad typological perspective (§ 7.1); next, we take an areal perspective, addressing the question to what extent the AP systems are similar to those of the surrounding Austronesian languages, and suggest where contact could have played a role in shaping the numerals (§ 7.2).

### 7.1 Typological rarities in AP numeral(s)

In Table 11 we summarise the various systems that AP languages use to form cardinals ‘five’ through ‘ten’. The Arabic numerals represent the numeral morphemes used in compounds, and the English words represent the lexical items that combine with those formatives. Thus, a compound numeral like Wersing Kolana *wetiηnuη* ‘six’ would be transcribed as ‘5 single’ as it is made up of *wetiη*, the morpheme for ‘5’, and a lexeme *nuη*, meaning ‘single’, while a compound numeral like Teiwa *jesraq* ‘seven’ would be transcribed as ‘5 2’ as it is a com-

Table 11: Morpheme patterns in AP cardinals ‘five’ through ‘ten’

	‘five’	‘six’	‘seven’	‘eight’	‘nine’	‘ten’
Northern Pantar	5	6	5 2	5 3	5 4	10 1
Central Alor	5	6	5 2	5 3	5 4	10 (1)
East Alor	5	5 1	5 2	5 3	5 4	10
Wersing Kolana	5	5 single	5 2	5 3	5 4	10
Kui	5	6	5 2	[2] 4	5 4	10 1
Straits-West-Alor	5	6	7 3	[10] less 2	[10] less 1	10 1
Western Pantar	5	5 single	7 2	7 3	1 take away [10]	10 1

pound of morphemes for ‘5’ and ‘2’. Square brackets ‘[]’ represent absent surface elements that are assumed to be part of the earlier numeral construction as we reconstructed it, while round brackets ‘()’ represent elements which may or may not be present depending on the details of the language in question.

Two major typological points are of interest in the AP numeral systems. The first is that they combine a mono-morphemic ‘six’ with base-five compounds for numerals ‘seven’ to ‘nine’. We reconstruct this system for proto-AP, and it is presently reflected in languages of northern Pantar and central Alor. Cross-linguistically, this is a rather uncommon numeral system; it is much more common to have a system where the quinary base is used in forming all numerals ‘six’ through nine’,<sup>15</sup> as attested in the languages of east Alor. However, because pAP \*talam ‘six’ is a reflex of the higher pTAP \*talam ‘six’ (Schapper, Huber & van Engelenhoven this volume), this form must be considered older than the quinary numeral for ‘six’. In other words, where we find [5 1] ‘six’, this is viewed as a later extension of the quinary system that was already in use for ‘seven’ through ‘nine’ in proto-AP.<sup>16</sup>

A second point of typological interest are the subtractive decimal systems used in the Straits-West-Alor languages, where numerals ‘seven’ through ‘nine’ are formed by subtraction,<sup>17</sup> while ‘six’ is monomorphemic. Systems like this, where subtraction is used in the formation of more than one numeral, and where such subtractive forms occur *alongside* a monomorphemic form for ‘six’, are crosslinguistically uncommon.

## 7.2 AP numerals in their areal context

It is useful to complement the genealogical perspective of sections 4 to 6 with an areal perspective, and compare the numeral system patterns in the AP languages with those of the Austronesian languages in their immediate vicinity, to see what this might tell us about the history of AP numerals. Where similar forms or patterns are found, we may ask whether there is evidence that these are contact-induced. In this section, we look at the evidence that may suggest influence from AP languages into nearby Austronesian languages, followed by the evidence suggesting influence in the opposite direction. We also point out cases where the data currently available are inconclusive.

<sup>15</sup> Harald Hammarström, p.c. 2012, based on his extensive numeral database reported on in Hammarström (2010).

<sup>16</sup> This contrasts with the view expressed by Vatter, who considered monomorphemic ‘six’ to be a ‘deviation’ (‘Abweichung’) from the base-five compounds ‘six’ (1932: 279-280).

<sup>17</sup> This is independent of the substitution of the Austroneisan base into ‘seven’.

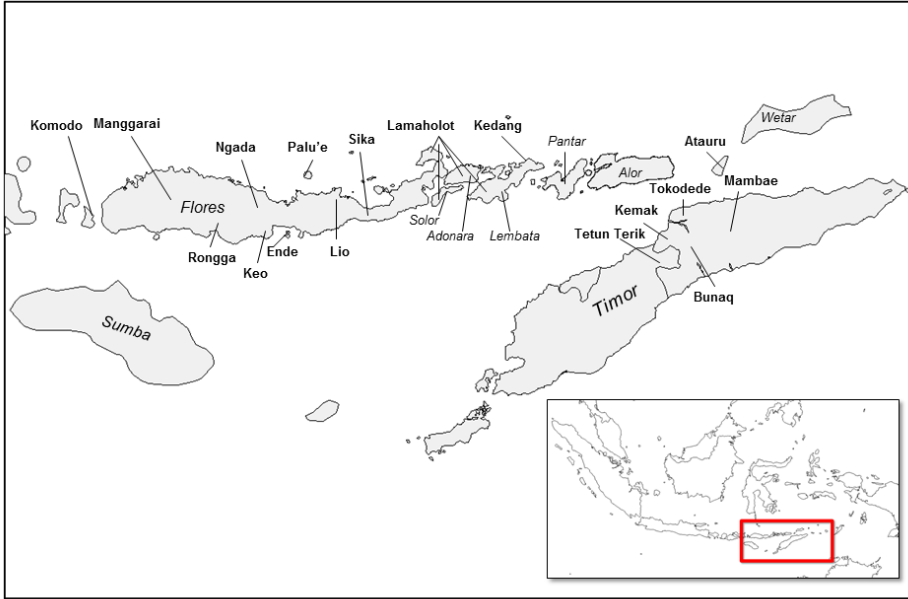


Figure 1: Austronesian (Austronesian) languages to the west and south of Alor-Pantar. Names in bold are language names, names in *italics* are names of islands.

It is a well-established fact that proto-Austronesian (pAN) had a decimal system, with numerals ‘one’ through ‘nine’ all being simple mono-morphemic words. Blust (2009: 268) claims that outside of Melanesia few Austronesian languages have innovated complex - additive, subtractive or multiplicative - numerals for ‘one’ to ‘ten’. The Austronesian languages around AP, however, show a notable clustering of just such innovations. We compiled numeral data for 32 Austronesian languages spoken west and south-east of Alor and Pantar (see Appendix A.2 and A.3). In these, we observe three distinct patterns of innovations in the formation of numerals ‘six’ through ‘nine’, reflected in nine modern Austronesian languages (Table 13). These are: the Timor pattern (1-5, 5+1, 5+2, 5+3, 5+4, 10), the Lembata pattern (1-7, 4x2, 5+4, 10), and the Flores pattern (1-5, 5+1, 5+2, 2x4, 10-1, 10). Proto-Austronesian numerals are provided for comparative purposes in the top row.

Table 12: Mixed numeral systems in proto-Austronesian and the Austronesian languages of Flores, Lembata and Timor (1-5)

		‘one’	‘two’	‘three’	‘four’	‘five’
	pAN	* <i>esa</i> ~ * <i>isa</i>	* <i>duSa</i>	* <i>telu</i>	* <i>Sepat</i>	* <i>lima</i>
Flores	Rongga	( <i>e</i> ) <i>sa</i>	<i>ɭua</i>	<i>telu</i>	<i>wutu</i>	<i>lima</i>
	Ende	<i>sa</i>	<i>zua</i>	<i>tela</i>	<i>wutu</i>	<i>lima</i>
	Ngadha	<i>esa</i>	<i>zua</i>	<i>telu</i>	<i>vutu</i>	<i>lima</i>
	Nage	<i>esa</i>	<i>dua</i>	<i>telu</i>	<i>wutu</i>	<i>lima</i>
	Kéo †	<i>ha?esa</i>	<i>?esa rua</i>	<i>?esa tedu</i>	<i>?esa wutu</i>	<i>?esa dima</i>
	Lio	<i>asa</i>	<i>rua</i>	<i>təlu</i>	<i>sutu</i>	<i>lima</i>
Lembata	Kedang*	> <i>ude?</i>	<i>sue</i>	<i>tælu</i>	> <i>apa?</i>	<i>leme</i>
Timor	Mambae	<i>id</i>	<i>ru</i>	<i>teul</i>	<i>fat</i>	<i>lim</i>
	Tokodede	<i>iso</i>	<i>ru</i>	<i>telo</i>	<i>pat</i>	<i>lim</i>

† Kéo numerals appear with the default classifier *?esa* and/or the prefix *ha* ‘one’.

\* In Kedang orthography />/ preceding a vowel encodes that vowel as breathy (Samely 1991).

Innovative quinary numerals are found in the Austronesian languages across the three innovative types. In the north-central Timor languages, Tokodede and Mambae, we have quinary numerals for numerals from ‘six’ through ‘nine’, a pattern that stands out against the typically conservative numerals systems of the Austronesian languages elsewhere on Timor (Naueti being an exception, see Schapper & Hammarström (2013) on the possible reasons for the quinary numerals in Naueti). It is notable that the close inland relative of Tokodede and Mambae, Kemak, has no base-5 numerals (see Appendix A.3). The appearance of this pattern in these languages may be result of contact with speakers of AP languages spoken on the south and east coast of Alor, such as Kula, Sawila and Wersing, located just a short sea crossing from the north of Timor. There is some linguistic evidence that contacts between these Alor groups and those of north-central Timor existed: the additive operators in the central-east Alor languages Sawila, Kula and Wersing (see Table 9) seem to be borrowed from Tokodede (§ 6.2). In addition, oral traditions record contacts between groups in south-east Alor and north Timor. For instance, eastern Alor groups almost invariably trace their origins to pre-historic migrations from Timor (Wellfelt & Schapper 2013, Wellfelt pers. comm. 2013). Similarly, many songs in central-east Alor are sung in the Tokodede language and mention place names such as Likusaen and Maubara, which are located in the north of Timor in the area where Tokodede is spoken

Table 13: Mixed numeral systems in proto-Austronesian and the Austronesian languages of Flores, Lembata and Timor (6-10)

		‘six’	‘seven’	‘eight’
	pAN	<i>*enem</i>	<i>*pitu</i>	<i>*walu</i>
Flores	Rongga	<i>lima esa</i> 5 1	<i>limaua</i> 5 2	<i>luambutu</i> 2 4
	Ende	<i>limasa</i> 5 1	<i>limazua</i> 5 2	<i>ruabutu</i> 2 4
	Ngadha	<i>lima esa</i> 5 1	<i>limarua</i> 5 2	<i>ruabutu</i> 2 4
	Nage	<i>lima esa</i> 5 1	<i>lima zua</i> 5 2	<i>zua butu</i> 2 4
	Kéo †	<i>?esa dima ?esa</i> 5 1	<i>?esa dima rua</i> 5 2	<i>?esa rua mbutu</i> 2 4
	Lio	<i>lima əsa</i>	<i>lima rua</i>	<i>rua mbutu</i> 2 4
	Lembata Kedang*	<i>&gt;ænæŋg</i>	<i>pitu</i>	<i>butu rai</i> 4 2?
Timor	Mambae	<i>limnai nide</i> 5 1	<i>limnai rua</i> 5 2	<i>limnai telu</i> 5 3
	Tokodede	<i>wouniso</i> [5] 1	<i>wouru</i> [5] 2	<i>woutelo</i> [5] 3

		‘nine’	‘ten’
	pAN	<i>*siwa</i>	<i>*puluq</i>
Flores	Rongga	<i>taraesa</i> [10] 1	<i>sambulu</i> 1 10
	Ende	<i>trasa</i> [10] 1	<i>sabulu</i> 1 10
	Ngadha	<i>teresa</i> [10] 1	<i>habulu</i> 1 10
	Nage	<i>tea esa</i> [10] 1	<i>sa bulu</i> 1 10
	Kéo †	<i>?esa tera ?esa</i> [1 10] 1	<i>hambudu</i> 1 10
	Lio	<i>təra əsa</i> [10] 1	<i>sambulu</i> 1 10
	Lembata Kedang*	<i>leme &gt;apa?</i> 5 4	<i>pulu</i> 1 10
Timor	Mambae	<i>limnai pata</i> 5 4	<i>sikul</i>
	Tokodede	<i>woupat</i> [5] 4	<i>sagulu</i> 1 10

† Kéo numerals appear with the default classifier *?esa* and/or the prefix *ha* ‘one’.

\* In Kedang orthography />/ preceding a vowel encodes that vowel as breathy (Samely 1991).

(Wellfelt & Schapper 2013). However, as Wellfelt & Schapper (2013) argue, the directionality of the influence in the contact relations retrievable from such oral traditions and linguistic evidence is firmly flowing from Timor to Alor. The borrowing of quinary numerals from AP into Timor languages thus would appear to go against the other borrowing patterns, including that seen thus far in numerals. As such, whilst Alorese quinary numerals are the only such systems that are in contact with the Tokodede and Mambae and seem the best candidate for the innovative numeral formation, it remains to be explained why this pattern was able to spread to the Timor languages, when in oral traditions and language it is the Timor groups that are the source of influence on Alor and not a recipient of it.

The origin of the base-five numerals in the central-eastern Flores languages Rongga, Ende, Ngadha, Nage, Kéo, and Lio is yet more obscure. There is mounting evidence of a non-Austronesian substrate in the Austronesian languages of the Flores region (see, e.g., Capell 1976, Klamer 2012). Accordingly, we may hypothesize that the quinary forms of the Flores languages reflect a prehistoric Papuan (or non-Austronesian) substrate that had a quinary system for the lower cardinals. However, we currently lack any evidence to link the languages forming the substrate in the central-eastern Flores region to the AP languages as we know it today – the Flores substrate could just as well been part of a different non-Austronesian group.

For Kedang on north Lembata, however, we are on a firmer ground to say that it formed its numeral ‘nine’ on the basis of the quinary patterns used for ‘six’ through ‘nine’ in the AP languages of northern Pantar, which is located just east of the Kedang speaking area on Lembata (see Figure 1). Note that the Lamaholot dialects spoken around Kedang in south and west Lembata all lack quinary numerals (see Appendix A.2) so that Kedang ‘nine’ stands out as being different from its immediate Austronesian neighbours. In his ethnographic study of the Kedang, Barnes (1974) noted that the Kedang speakers are culturally very different from the Lamaholot groups on Lembata, instead showing cultural similarities with the AP groups of Alor and Pantar. For instance, unlike the Lamaholot, the Kedang are known for ‘the number of gongs [they] own and especially in the fact that [these] are used as bridewealth’ (Barnes 1974: 15), which is also a common practice in AP groups. The unique quinary form of Kedang ‘nine’ may well be a trace of cultural contact between Kedang and AP speakers on Pantar, for instance in bridewealth negotiations involving gongs.

In turn, we now investigate to what extent the numeral systems in AP languages have been influenced by nearby Austronesian languages. In sections 4-6,



we saw that some AP languages employ numerals containing morphemes that have been borrowed from Austronesian languages, in the following five contexts:

1. A reflex of pAN \*pitu was borrowed into the Straits-West Alor languages as a base in the numeral ‘seven’.
2. The Western Pantar numeral ‘six’ *hisnakkung* has an initial element *his-* that is a likely Austronesian borrowing (< PAN *esa~\*isa* ‘one’), and *hisnakkung* represents a partial calque of the [5 1] pattern found in Austronesian languages of Flores.
3. An additive operator with the approximate form /geresin/ was borrowed into east Alor languages from Tokodede (north Timor).
4. Reflex(es) of pAN \*Ratus ‘hundred’ were borrowed from the Flores-Lembata Austronesian languages into the languages of Pantar, and to a lesser extent Alor.
5. Reflex(es) of pAN \*libu ‘thousand’ were borrowed from Flores-Lembata Austronesian languages into AP languages across the board.

The pattern for Kui *tadusa* ‘eight’ seems to be formed a multiplicative pattern 2x4 due to the second element appearing to be derived from *usa* ‘four’. This multiplicative pattern is otherwise unknown in AP languages, but is found in the Austronesian languages of central-eastern Flores (Ende, Lio, Ngadha, Rongga, Keo). Whilst the Kui are today not directly adjacent to any of the Austronesian languages of Flores with multiplicative ‘eight’, there are indications that they may have had fairly intensive contact with Austronesian speakers from the west. Kui oral tradition holds that the royal family of the group migrated to Alor from Flores (Emilie Wellfelt pers. comm.). Hägerdal (2012:38, fn. 36) cites evidence that the Kui were part of a league consisting of the five princedoms Pandai, Baranusa, Blagar, Alorese and Kui. Today, Pandai, Baranusa and Alor are locations where Alorese is spoken, an Austronesian language closely related to Lamaholot in the Flores region (Klamer 2011, 2012). In the historical period, the Kui king is also widely recorded to have owned boats running trade routes between Alor and Kupang in West Timor and islands of the Solor archipelago (Emilie Wellfelt pers. comm., Hägerdal 2012). It is therefore possible that the Kui were once in close contact with speakers of (an) Austronesian language(s) from the Flores region, and that this contact might be the ultimate source of their base-4 numeral ‘eight’.

Finally, recall that forming ‘nine’ by subtraction ([10]-1) is found in the AP languages of Straits-West-Alor, while in the Austronesian languages of central-east

Flores, monomorphemic ‘nine’ (proto-Austronesian \*siwa) has been replaced with a subtractive compound containing two formatives: a reflex of proto-Austronesian \*esa ‘one’, and an unanalysable initial element (\*tar). There is no obvious explanation how the subtractive ‘nine’ entered this group of Flores languages. Neither can we explain the origin of the subtractive pattern in proto-Straits-West-Alor. We have argued that in this proto-language, subtractive ‘nine’ replaced the original pAP base-five form of ‘nine’ [5 4], and involved reflexes of pAP \*nuk ‘one’, subsequently extending the subtractive system to ‘eight’ and ‘seven’. So it is the subtractive pattern that is similar across the Flores and Straits-West Alor groups, not the lexemes themselves. The geographical closeness of the groups, combined with the relative rarity of subtractive systems in both Austronesian and Alor-Pantar languages, may be suggestive of a (possibly ancient) structural diffusion. On the other hand, we cannot exclude the possibility that the forms were innovated independently in both groups: as Schapper & Hammarström (2013) point out, innovation of subtractive numerals has occurred independently in proto-Malay, central Maluku and south-east Sulawesi.

In short, contact-induced borrowings of both forms and structures (‘matter’ and ‘patterns’) have played a role in shaping some of the numerals in the Alor-Pantar and nearby Austronesian languages. Some contacts took place in historical times, and are supported by historical and ethnographic data, others are likely to be of more ancient date and must remain hypothetical. There are also similarities that cannot be traced back to contact.

## **8 Conclusions and discussion**

From this comparative study of the numeral paradigms in 19 AP language varieties we draw three types of conclusions: (i) about the morphological make up of the numeral compounds; (ii) about typological rarities in the AP numeral systems, and (iii) about the subgrouping and history of the AP language group.

Morphologically, AP cardinals above ‘six’ consist of minimally two formatives. Additive base-five forms involve two (reflexes of) numerals and no marker for addition. Subtractive base-ten forms involve a numeral and an unanalysable initial element that appears to go back to a morpheme originally meaning something like ‘less’ or ‘take away’. With one exception, the numerals ‘ten’ are compounds of ‘ten’ and ‘one’, and the decades are formed accordingly. Numerals ‘one hundred’ and ‘one thousand’ are structured in the same way, expressing multiplication of the base with juxtaposed numerals in which the highest numeral precedes the lowest. Numerals in between decades are expressed as phrases, involving an

additive operator (proto-AP \*wali(ŋ) ‘add, (do) again’).

Typologically, the constellations of numerals ‘six’ through ‘nine’ in AP represent two rare patterns. The first rarity is the combination of a mono-morphemic ‘six’ with quinary forms ‘seven’ through ‘nine’ found in many languages across the two islands, and reconstructed to proto-AP. The second rarity is the occurrence of subtractive base-ten systems alongside a monomorphemic ‘six’ as found in the Straits-West-Alor languages. Typologically interesting are the mathematically ‘incorrect’ numerals found in some of the languages: a ‘seven’ that mathematically should be ‘four’ (Straits-West Alor), a ‘seven’ and ‘eight’ that mathematically should be ‘nine’ and ‘ten’ (Western Pantar), and an ‘eight’ that would literally translate as ‘minus four’ in Kui. These forms all arose through reanalysis of the numeral value of the base as different from its etymological source. Finally, it is of typological interest to consider the non-numeral lexemes that are incorporated into AP numerals: the (ad)verbs ‘less’ and ‘take away’ as part of subtractive numerals; *mi-*, an originally locative morpheme deriving decades; and the word ‘single’ standing in for the numeral ‘one’ in compound numerals for ‘six’.

Historically, this study has provided information on the numeral system of proto-AP, and additional details on affiliations and distinctions between members of the AP group that may be used as evidence to construct particular subgroups within the family. The proto-AP numeral system was a mixed quinary and decimal system, with a monomorphemic ‘six’ (i.e. 1, 2, 3, 4, 5, 6, [5 2], [5 3], [5 4], [10 1]). The arithmetic operations involved were addition and multiplication. Over time, the system was complicated by reorganisations of patterns as well as borrowings of numeral bases, or patterns, or both. As a result, some modern languages have introduced subtractive procedures instead of, or along with, addition and multiplication. Some languages incorporated non-numeral formatives into their numerals.

Numeral forms were reconstructed to different nodes in the the AP family, as summarised in Table 14. The table is to be read from left to right. The left-most column represents the oldest numeral forms, that is, those that can be reconstructed to proto-AP. Numerals ‘one’ through ‘six’ in this proto-language were monomorphemic forms, while ‘seven’ through ‘nine’ were regular quinary forms. The right-hand columns represent numeral innovations which can be reconstructed to different subgroups of the AP family.<sup>18</sup>

<sup>18</sup> The ordering of right-hand columns is by number of languages; it should not be interpreted as representing a chronology of the age of subgroups in the case of proto-Straits-West-Alor (pSWA) and proto-Pantar (pP). Naturally, proto-Central-East-Alor (pCEA), proto-East Alor (pEA) and proto-East Alor Montane (pEAM) can be taken to represent a chronological se-

Table 14: Numeral (pattern) reconstructions for ‘one’ through ‘ten’ in AP subgroups

	Proto-Alor- Pantar	Proto-Straits-West- Pantar	Proto- Pantar	Proto-Central- East-Alor	Proto-East Alor	Proto-East Alor	Mon- tane
‘one’	* <i>nuk</i>						* <i>sundana</i>
‘two’	* <i>araq</i>						
‘three’	* <i>(a)tiga</i>						
‘four’	* <i>buta</i>						
‘five’	* <i>yiwesin</i>						
‘six’	* <i>talam</i>					* <i>arasiku</i>	
‘seven’	5 2	* <i>butitoga</i>	* <i>yewasraqo</i>				
‘eight’	5 3	* <i>turarok</i>	* <i>yesantig</i>				
‘nine’	5 4	* <i>tukarinuk</i>	* <i>yesanut</i>				
‘ten’	* <i>qar</i>			* <i>adayaku</i>			

Translated into a tree, the reconstruction of numerals in AP languages yields the structure in Figure 2.

We see that that there are patterns and forms found in the Pantar languages (except Western Pantar) are clearly separate from those in the languages of Alor; the Straits West Alor languages (Blagar, Reta, Kabola, Adang, Hamap, Klon) share patterns and forms amongst themselves that are not shared with other AP languages; and we argued the same to be the case for the languages of central and east Alor. The subgrouping membership of the Kui and Western Pantar is problematic; their grouping within pSWA is tentative and rests on their possessing some innovative morphemes (i.e., reflexes of \**buti-* and \**tukari*) in common with the main Straits-West Alor languages proper, though with different functions in Kui and Western Pantar.

The preliminary reconstruction of Proto-AP based on sound changes as reported in Holton et al. (2012) and (Holton & Robinson this volume) focuses on showing the relatedness of all AP languages. Little work has been done on the sound changes defining lower-level subgroups of AP languages. Nevertheless, there are some correspondences that can be observed. For instance, the pSWA subgroup we define (without the problematic inclusion of Kui and Western Pantar) is also supported by the sound change \**s* > *h*. Further study of lower level sound changes is needed to test whether all the subgroups we posit here on the basis of the morphological evidence of numerals are valid.

In sum, cardinal numerals in the Alor-Pantar languages are fertile ground for understanding how diverse numeral systems can evolve in related languages. In particular, Alor-Pantar languages provide us with unique, typological insights into the historical changes and influences that can complicate and prompt reorganisations of patterns of numeral formation and borrowings into the numeral paradigm.

---

quence, since pEA forms a subgroup of pCEA, and pEAM a subgroup of pEA.

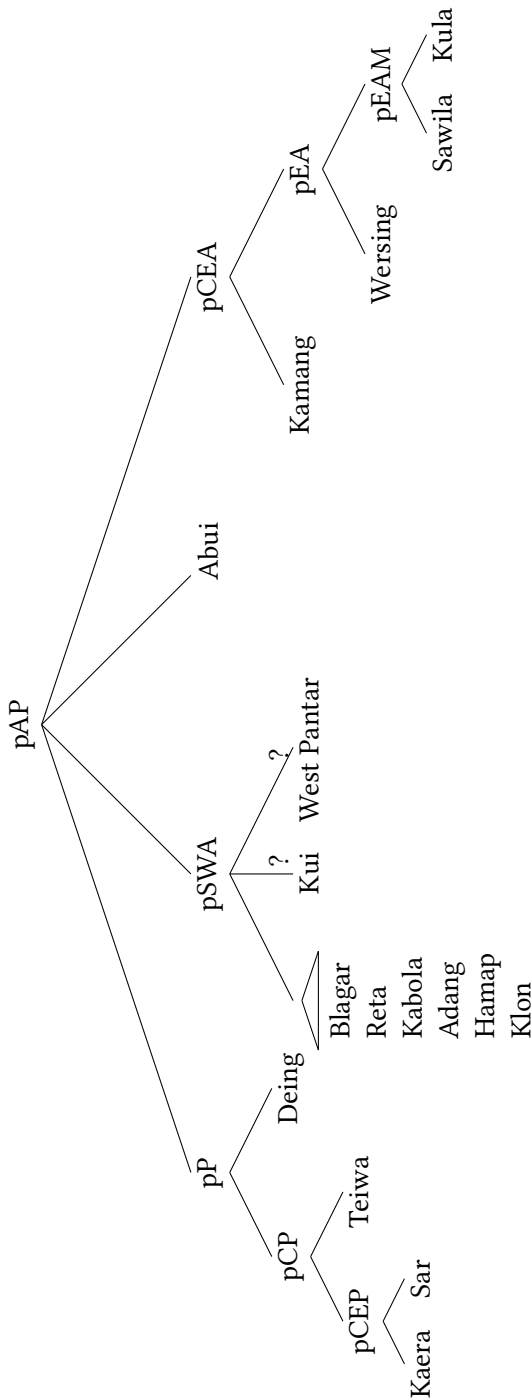


Figure 2: Tree of the AP languages based on numeral innovations

## 9 Sources

Sources of the language data cited in the text and the Appendices are given in the table below. We provide information about the dialect in cases where unpublished sources are used, or where multiple dialects are cited.

Abui (AP)	Kratochvíl (2007), Schapper fieldnotes 2010
Adang (AP, Pitungbang dialect)	Robinson fieldnotes 2010
Alorese (AN)	Klamer (2011)
Amarasi (AN)	Bani & Grimes (2011)
Atauro (AN)	Schapper fieldnotes 2007
Blagar (AP, Bama dialect)	Robinson fieldnotes 2010
Blagar (AP, Dolabang dialect)	Hein Steinhauer p.c. 2011
Bunaq (TAP, Lamaknen)	Schapper (2010)
Bunaq (TAP, Manufahi)	Schapper fieldnotes 2007
Dadu'a (a.k.a. Galoli) (AN)	Penn (2006)
Dhao	Grimes, Ranoh & Aplugi (2008)
Deing (AP)	B. Volk fieldnotes 2008
Ende (AN)	Aoki & Nakagawa (1993)
Hamap (AP)	Baird fieldnotes 2003
Idate (AN)	Klamer fieldnotes 2002
Ilongot (AN)	ABVD
Kabola (AP)	Robinson fieldnotes 2010
Kaera (AP)	Klamer fieldnotes 2005
Kamang (AP)	Schapper fieldnotes 2010, 2011
Kedang (AN)	Samely (1991)
Kemak (AN, Atabai dialect)	Klamer fieldnotes 2002
Kéo (AN)	Baird (2002)
Klon (AP)	Baird (2008)
Komodo (AN)	Verheijen (1982)
Kui (AP)	Baird fieldnotes 2003, Holton fieldnotes 2010

Kula (AP)	Holton fieldnotes 2010, Nicholas Williams p.c. 2011
Lakalei (AN)	Klamer fieldnotes 2002
Lamaholot (AN, Lewoingu dialect)	Nishiyama & Kelen (2007)
Lamaholot (AN, Lewotobi dialect)	Naonori Nagaya p.c. 2011
Lamaholot (AN, Lewolema dialect)	Pampus (2001)
Lamaholot (AN, Solor dialect)	Klamer fieldnotes 2002
Lamaholot (AN, Adonara)	Philippe Grangé p.c. 2011
Lamaholot (AN, Lamalera dialect)	Keraf (1978)
Lio (AN)	Sawardo, Tarno & Kusharyanto (1987: 127-137, 44, 57, 60, 75, 110), Arndt (1933)
Mambae (AN, Ainaro dialect)	Schapper fieldnotes 2007
Manggarai (AN)	Verheijen (1967:518); Verheijen (1970:173)
Nage (AN)	Gregory Forth p.c. 2011
Ngadha (AN)	Arndt (1961)
Palu'e (AN)	ABVD
Reta (AP)	Robinson fieldnotes 2010
Rembong (AN)	Verheijen (1978)
Rongga (AN)	Arka et al. (2007)
Sar (AP)	Baird fieldnotes 2003; Robinson fieldnotes 2010
Sika (AN)	Pareira & Lewis (1998), Calon (1890)
Teiwa (AP)	Klamer (2010)
Tetun Fehan (AN)	Van Klinken (1999:100)
Tokodede (AN, Licissa dialect)	Schapper fieldnotes 2007
Uab Meto (AN)	Middelkoop (1950:421-424)
Ujir (AN)	Schapper fieldnotes
Waima'a (AN)	Hull (2002)
Western Pantar (AP)	Holton (nd)
Wersing (AP)	Holton fieldnotes 2010, Schapper & Hendery (2014)



## A Appendix

### A.1 Cardinal numerals in the Alor-Pantar languages

Varieties within a language are indicated by the name of one of the places where the dialect is spoken, though often dialects cover more than one place.

Table 15: Numerals ‘one’ through ‘four’

Location	Language	‘one’	‘two’	‘three’	‘four’
Pantar	Western Pantar	<i>anuku</i>	<i>alaku</i>	<i>atiga</i>	<i>atu</i>
	Deing	<i>nuk</i>	<i>raq</i>	<i>atig</i>	<i>ut</i>
	Sar	<i>nuk</i>	<i>raq</i>	<i>tig</i>	<i>ut</i>
	Teiwa	<i>nuk</i>	<i>(ha)raq</i>	<i>jerig</i>	<i>ut</i>
	Kaera	<i>nuk(u)</i>	<i>(a)rax-</i>	<i>(i/u)tug</i>	<i>ut</i>
Straits	Blagar-Bama	<i>nuku</i>	<i>akur</i>	<i>tuge</i>	<i>ut</i>
	Blagar-Dolabang	<i>nu</i>	<i>aru</i>	<i>tue</i>	<i>buta</i>
	Reta	<i>anu</i>	<i>alo</i>	<i>atoga</i>	<i>w/buta</i>
W Alor	Kabola	<i>nu</i>	<i>olo</i>	<i>towo</i>	<i>ut</i>
	Adang	<i>nu</i>	<i>alo</i>	<i>tuo</i>	<i>ut</i>
	Hamap	<i>nu</i>	<i>alo</i>	<i>tof</i>	<i>ut</i>
	Klon	<i>nuk</i>	<i>orok</i>	<i>tonj</i>	<i>ut</i>
	Kui	<i>nuku</i>	<i>oruku</i>	<i>siwa</i>	<i>usa</i>
C&E Alor	Abui	<i>nuku</i>	<i>ajoku</i>	<i>sua</i>	<i>buti</i>
	Kamang (Atoitaa)	<i>nok</i>	<i>ok</i>	<i>su</i>	<i>biat</i>
	Kamang (Takailubui)	<i>nok</i>	<i>ok</i>	<i>su</i>	<i>biat</i>
	Sawila	<i>sundana</i>	<i>jaku</i>	<i>tuo</i>	<i>ara:si:ku</i>
	Kula	<i>sona</i>	<i>jakwu</i>	<i>tu</i>	<i>arasiku</i>
	Wersing	<i>no</i>	<i>joku</i>	<i>tu</i>	<i>arasoku</i>

Table 16: Numerals ‘five’ through ‘nine’

Location	Language	‘five’	‘six’	‘seven’	‘eight’	‘nine’
Pantar	Western Pantar	<i>jasin</i>	<i>hisnakkun</i>	<i>betalaku</i>	<i>betiga</i>	<i>anukutannan</i>
	Deing	<i>asan</i>	<i>talan</i>	<i>jewasrak</i>	<i>santig</i>	<i>sanut</i>
	Sar	<i>awan</i>	<i>tejan</i>	<i>jisraq</i>	<i>jinatig</i>	<i>jinaut</i>
	Teiwa	<i>jusan</i>	<i>tia:m</i>	<i>jesraq</i>	<i>jesnerig</i>	<i>jesna?ut</i>
Straits	Kaera	<i>isim</i>	<i>tia:m</i>	<i>jesrax-</i>	<i>jentug</i>	<i>jeniut</i>
	Blagar-Bama	<i>isin</i>	<i>tajan</i>	<i>titu</i>	<i>tuakur</i>	<i>tukurunuku</i>
	Blagar-Dolabang	<i>isin</i>	<i>talin</i>	<i>bititu</i>	<i>tuaru</i>	<i>turinu</i>
	Reta	<i>avehan</i>	<i>talaun</i>	<i>biitoga</i>	<i>tulalo</i>	<i>tukanu</i>
W Alor	Kabola	<i>iwesin</i>	<i>talan</i>	<i>wuitito</i>	<i>turlo</i>	<i>ti?inu</i>
	Adang	<i>ifihin</i>	<i>talan</i>	<i>ititɔ</i>	<i>turlo</i>	<i>ti?enu</i>
	Hamap	<i>ivehin</i>	<i>talan</i>	<i>itito</i>	<i>turalo</i>	<i>tieu</i>
	Klon	<i>eweh</i>	<i>tlan</i>	<i>uson</i>	<i>tidorok</i>	<i>tukainuk</i>
C & E Alor	Kui	<i>jesan</i>	<i>talama</i>	<i>jesaroku</i>	<i>tadusa</i>	<i>jesanusa</i>
	Abui	<i>jetin</i>	<i>tala:ma</i>	<i>jetinajoku</i>	<i>jetinsua</i>	<i>jetinbuti</i>
	Kamang (Takailubui)	<i>wesin</i>	<i>ta:ma</i>	<i>wesinok</i>	<i>wesinsu</i>	<i>wesinbiat</i>
	Kamang (Atoitaa)	<i>iwesin</i>	<i>isinok</i>	<i>isinok</i>	<i>isinsu</i>	<i>isinbiat</i>
	Sawila	<i>jo:tin</i>	<i>jo:tingsundana</i>	<i>jo:tinjaku</i>	<i>jo:tingtuo</i>	<i>jo:tinara:si:ku</i>
	Kula	<i>jawatena</i>	<i>jawatensona</i>	<i>jawatenjak-wu</i>	<i>jawatentu</i>	<i>jawatenarasiku</i>
	Wersing	<i>wetin</i>	<i>wetinun</i>	<i>wetinjoku</i>	<i>wetintu</i>	<i>wetinarasoku</i>

Table 17: Numerals ‘ten’ and the formation of decades

		‘ten’	‘twenty’	‘thirty’
Pantar	Western Pantar	<i>ke anuku</i>	<i>ke alaku</i>	<i>ke atiga</i>
	Deing	<i>qar nuk</i>	<i>qar raq</i>	<i>qar atig</i>
	Sar	<i>qar nuk</i>	<i>qar raq</i>	<i>qar tig</i>
	Teiwa	<i>qa:r nuk</i>	<i>qa:r raq</i>	<i>qa:r jerig</i>
	Kaera	<i>xar nuko</i>	<i>xar raxo</i>	<i>xar tug</i>
Straits	Blagar-Bama	<i>qar nuku</i>	<i>qar akur</i>	<i>qar tuge</i>
	Blagar-Dolabang	<i>?ari nu</i>	<i>?ari aru</i>	<i>?ari tue</i>
	Reta	<i>kara nu</i>	<i>kara alo</i>	<i>kara atoga</i>
West Alor	Kabola	<i>kar nu</i>	<i>kar ho(?)olo</i>	<i>kar towo</i>
	Adang	<i>?er nu</i>	<i>?er alo</i>	<i>?er tuo</i>
	Hamap	<i>air nu</i>	<i>air alo</i>	<i>air tof</i>
	Klon	<i>kar nuk</i>	<i>kar orok</i>	<i>kar toŋ</i>
	Kui	<i>kar nuku</i>	<i>kar oruku</i>	<i>kar siwa</i>
C & E Alor	Abui	<i>kar nuku</i>	<i>kar ajoku</i>	<i>kar sua</i>
	Kamang	<i>ata:k nok</i>	<i>ata:k ok</i>	<i>ata:k su</i>
	Sawila	<i>ada:ku</i>	<i>ada:ku maraku</i>	<i>ada:ku matua</i>
	Kula	<i>adajakwu</i>	<i>mijakwu</i>	<i>mitua</i>
	Wersing	<i>adajoku</i>	<i>adajoku mijoku</i>	<i>adajoku mitu</i>

Table 18: Numerals with bases ‘100’ and ‘1000’

		‘100’	‘200’	‘1000’	‘2000’
Pantar	West Pantar	<i>ratu</i>	<i>ratu alaku</i>	<i>(a)ribu (ye)</i>	<i>ribu (alaku)</i>
	Deing	<i>aratu nuk</i>	<i>aratu raq</i>	<i>aribu nuk</i>	<i>aribu raq</i>
	Sar	<i>ratu nuk</i>	<i>ratu raq</i>	<i>ribu nuk</i>	<i>ribu raq</i>
	Teiwa	<i>ratu nuk</i>	<i>ratu (ha)raq</i>	<i>ribu nuk</i>	<i>ribu (ha)raq</i>
	Kaera	<i>ratu nuk</i>	<i>ratu rax-</i>	<i>ribu nuk</i>	<i>ribu rax-</i>
Straits	Blagar-Bama	<i>ratu nuk</i>	<i>ratu akur</i>	<i>ribu nuk</i>	<i>ribu akur</i>
	Blagar-Dolabang	<i>ratu nu</i>	<i>ratu aru</i>	<i>ribu nu</i>	<i>ribu aru</i>
	Reta	<i>ratu anu</i>	<i>ratu alo</i>	<i>ribu ano</i>	<i>ribu alo</i>
W Alor	Kabola	<i>rat nu</i>	<i>rat ho(?)olo</i>	<i>rib nu</i>	<i>rib ho(?)olo</i>
	Adang	<i>rat nu</i>	<i>rat alo</i>	<i>rib nu</i>	<i>rib alo</i>
	Hamap	<i>rat nu</i>	<i>rat alo</i>	— <sup>†</sup>	—
	Klon	<i>eska nok</i>	<i>eska orok</i>	—	—
C & E Alor	Kui	<i>asaga</i>	<i>asaga oruku</i>	<i>rab nuk</i>	<i>rab oruku</i>
	Abui	<i>aisaha nu</i>	<i>aisaha ajoku</i>	<i>rifi nuk</i>	<i>rifi ajoku</i>
	Kamang (A/T)	<i>asaka nok</i>	<i>asaka ok</i>	<i>libu nok</i>	<i>libu ok</i>
	Sawila	<i>asaka dana</i>	<i>asaka jaku</i>	<i>ri:bu dana</i>	<i>ri:bu jaku</i>
	Kula	<i>gasaka</i>	<i>gasaka jakwana</i>	<i>rib dena</i>	—
	Wersing	<i>aska</i>	<i>aska joku</i>	<i>ribu no</i>	—

† ‘—’ denotes that no data is available for this numeral

## A.2 Numerals ‘one’ to ‘ten’ in Austronesian languages W of Alor-Pantar

Table 19: Numerals ‘one’ to ‘five’ in Austronesian languages W of Alor-Pantar.

Location	Language	‘one’	‘two’	‘three’	‘four’	‘five’
	PAN	*esa~*isa	*duSa	*telu	*Sepat	*lima
Komodo	Komodo	sa, se-	rua	telu	pa?	lima
Flores	Manggarai	esa	sua	telu	pat	lima
	Rongga	(e)sa	ɿua	telu	wutu	lima
	Rembong	sa, sa?	zta	telu	pat	lima
	Ende	sa	zua	tela	wutu	lima
	Ngadha	esa	zua	telu	vutu	lima
	Nage	esa	dua	telu	wutu	lima
	Kéo†	ha?esa	?esa rua	?esa tedu	?esa wutu	?esa dima
	Lio	əsa	rua	təlu	sutu	lima
	Sika	ha	rua	telu	hutu	lima
	Palu’e	a	rua	təlu	ba	lima
	Lamaholot-Lewoingu	to?u	rua	təlo	pak	lema
	Lamaholot-Lewotobi	to?u	rua	təlo	pa	lema
	Lamaholot-Lewolema	to?u	rua	təlo	pat	lema
Solor	Lamaholot-Solor	to?u	rua	təlo	pa	lema
Adonara	Lamaholot-Adonara	to?u	rua	təlo	pat	lema
Lembata (Lomblen)	Lamaholot-Lamalera	tou	rua	telo	pa	lema
	Kedang†	>ude?	sue	tælu	>apa?	leme
	Alorese-Baranusa	to	rua	talau	pa	lema
Pantar	Alorese-Alor Kecil	tou	rua	telo	pa	lema

† Kéo numerals appear with the default classifier ?esa and/or the prefix ha ‘one’. In Kedang orthography />/ preceding a vowel encodes that vowel as breathy (Samely 1991)

Table 20: Numerals ‘six’ to ‘ten’ in Austronesian languages W of Alor-Pantar

Location	Language	‘six’	‘seven’	‘eight’	‘nine’	‘ten’
Komodo	PAN	*enem	*pitu	*walu	*siwa	*puluq
	Komodo	<i>nemu</i>	<i>pitu</i>	<i>walu</i>	<i>siwa</i>	<i>pulu</i> , <i>sampulu</i>
Flores	Manggarai	<i>enem</i>	<i>pitu</i>	<i>alo</i>	<i>ciok</i>	<i>pulu</i> <i>cempulu</i> <i>cepulu</i> <i>campulu</i>
	Rongga	<i>limaesa</i>	<i>limaua</i>	<i>luambutu</i>	<i>taraesa</i>	<i>sambulu</i>
	Rembong	<i>non</i>	<i>pitu?</i>	<i>walu?</i>	<i>siwa?</i>	<i>(se)puluh</i> <i>/pulu?</i>
	Ende	<i>limasa</i>	<i>limazua</i>	<i>ruabutu</i>	<i>trasa</i>	<i>sabulu</i>
	Ngadha	<i>limaesa</i>	<i>limarua</i>	<i>ruabutu</i>	<i>teresa</i>	<i>habulu</i>
	Nage	<i>lima esa</i>	<i>lima zua</i>	<i>zua butu</i>	<i>tea esa</i>	<i>sa bulu</i>
	Kéo†	<i>?esa dima ?esa</i>	<i>?esa dima rua</i>	<i>?esa rua mbutu</i>	<i>?esa tera ?esa</i>	<i>ha mbudu</i>
	Lio	<i>lima əsa</i>	<i>lima rua</i>	<i>rua mbutu</i>	<i>təra əsa</i>	<i>sambulu</i>
	Sika	<i>əna</i>	<i>pitu</i>	<i>walu</i>	<i>hiwa</i>	<i>pulu</i> <i>pulu ha</i>
	Palu’e	<i>əne</i>	<i>bitu</i>	<i>valu</i>	<i>iva</i>	<i>apulu</i>
	Lamaholot-	<i>nəmən</i>	<i>pito</i>	<i>buto</i>	<i>hiwa</i>	<i>pulo</i>
	Lewoingu					
	Lamaholot-	<i>namu</i>	<i>pito</i>	<i>buto</i>	<i>hiwa</i>	<i>pulo</i>
	Lewotobi					
	Lamaholot-	<i>nəm(ə)</i>	<i>pito</i>	<i>buto</i>	<i>hiwa</i>	<i>pulok</i>
	Lewolema					
	Solor	<i>nəmə</i>	<i>pito</i>	<i>wutu</i>	<i>hiwa</i>	<i>pulo?</i> <i>pulok</i>
	Adonara	<i>nəm(ə)</i>	<i>pito</i>	<i>buto</i>	<i>hiwa</i>	<i>pulo</i>
Lembata (Lomblen)	Lamaholot-	<i>nemu</i>	<i>pito</i>	<i>buto</i>	<i>hifa</i>	<i>pulo</i>
	Lamalera					
	Kedang†	<i>&gt;ænæŋ</i>	<i>pitu</i>	<i>buturai</i>	<i>leme&gt;apa?</i>	<i>pulu</i>
Pantar	Alorese-	<i>namu</i>	<i>pito</i>	<i>buto</i>	<i>hifa</i>	<i>karto</i>
	Baranusa					
	Alorese-	<i>nemu</i>	<i>pito</i>	<i>buto</i>	<i>hifa</i>	<i>kartou</i>
	Alor Kecil					

† Kéo numerals appear with the default classifier *?esa* and/or the prefix *ha* ‘one’. In Kedang orthography />/ preceding a vowel encodes that vowel as breathy (Samely 1991)

### A.3 Numerals ‘one’ to ‘ten’ in Austronesian languages S & E of Alor-Pantar

Table 21: Numerals ‘one’ to ‘five’ in Austronesian languages S & E of Alor- Pantar

Location	Language	‘one’	‘two’	‘three’	‘four’	‘five’
	PAN	*esa~*isa	*duSa	*telu	*Sepat	*lima
Rote	Dhao	<i>əf i</i>	<i>dua</i>	<i>təke</i>	<i>əpa</i>	<i>ləmi</i>
Atauro	Atauro	<i>hea</i>	<i>herua</i>	<i>hetelu</i>	<i>heat</i>	<i>helima</i>
Western Timor	Uab Meto ‡ <div>is this ‡ supposed to be here?</div>	<i>mese</i>	<i>nua</i>	<i>tenu</i>	<i>ha</i>	<i>nim</i>
	Amarasi	<i>es</i>	<i>nua</i>	<i>teun~tenu</i>	<i>ha:</i>	<i>ni:m~nima</i>
North-Central Timor	Mambae	<i>id</i>	<i>ru</i>	<i>teul</i>	<i>fat</i>	<i>lim</i>
	Tokodede	<i>iso</i>	<i>ru</i>	<i>telo</i>	<i>pat</i>	<i>lim</i>
Central Timor	Kemak	<i>sia</i>	<i>hurua</i>	<i>telu</i>	<i>pa:t</i>	<i>həlima</i>
	Lakalei	<i>isa</i>	<i>rua</i>	<i>telu</i>	<i>at</i>	<i>lima</i>
	Idate	<i>wisa</i>	<i>rua</i>	<i>telu</i>	<i>at</i>	<i>lima</i>
South-Central Timor	Tetun Fehan	<i>ida</i>	<i>rua</i>	<i>tolu</i>	<i>ha:t</i>	<i>lima</i>
North-Eastern Timor	Waima’a	<i>se</i>	<i>kairuo</i>	<i>kaitelu</i>	<i>kaiha:</i>	<i>kailime</i>
	Dadu’a	<i>isa</i>	<i>warua</i>	<i>watelu</i>	<i>wa:k</i>	<i>walima</i>

Table 22: Numerals ‘six’ to ‘ten’ in Austronesian languages S & E of Alor- Pantar

Location	Language	‘six’	‘seven’	‘eight’	‘nine’	‘ten’
	PAN	*enem	*pitu	*walu	*siwa	*puluq
Rote	Dhao	<i>əna</i>	<i>piɖa</i>	<i>aru</i>	<i>tʃeo</i>	<i>tʃaŋuru</i>
Atauro	Atauro	<i>henen</i>	<i>heitu</i>	<i>heau</i>	<i>hese</i>	<i>seŋulu</i>
Western Timor	Uab Meto	<i>ne</i>	<i>hitu</i>	<i>fanu‡</i>	<i>seo / sio</i>	<i>boʔes†</i>
	‡	<div style="border: 1px solid black; background-color: orange; padding: 5px; display: inline-block;"> is this ‡ supposed to be here? </div>				
	Amarasi	<i>nee</i>	<i>hiut~hitu</i>	<i>faun~fanu</i>	<i>seo / sea</i>	<i>boʔes</i>
North-Central Timor	Mambae	<i>limnainide</i>	<i>limnairua</i>	<i>limnaitelu</i>	<i>limnaipata</i>	<i>sikul</i>
	Tokodede	<i>wouniso</i>	<i>wouru</i>	<i>woutelo</i>	<i>woupat</i>	<i>sagulu</i>
Central Timor	Kemak	<i>hənem</i>	<i>hitu</i>	<i>balu</i>	<i>sibe</i>	<i>sapulu</i>
	Lakalei	<i>nen</i>	<i>hitu</i>	<i>walu</i>	<i>sia</i>	<i>sakulu</i>
	Idate	<i>nen</i>	<i>hitu</i>	<i>walu</i>	<i>sia</i>	<i>sanulu</i>
South-Central Timor	Tetun	<i>neen</i>	<i>hitu</i>	<i>walu</i>	<i>siwi</i>	<i>sanulu</i>
	Fehan					
North-Eastern Timor	Waima’a	<i>kainena</i>	<i>kaihitu</i>	<i>kaikaha</i>	<i>kaisiwe</i>	<i>base</i>
	Dadu’a	<i>wanee</i>	<i>waʔitu</i>	<i>waʔao</i>	<i>wasia</i>	<i>sanulu</i>

† *Boʔes* probably derives from *bua ès* ‘one collection’ according to Middelkoop (1950:421).

‡ *Fanu* ‘eight’ is used in the sense of ‘many’ “by reversing the last syllable” (i.e. as *faun*) (Middelkoop 1950: 422).



## Abbreviations

—	no data available	D	Blagar-D > Blagar-Dolabang
~REDUP	reduplication	E	Eastern
A	refers to the most agent-like argument of a canonical transitive verb	pAN	proto-Austronesian
		pAP	proto-Alor Pantar
		pCEA	proto-Central East Alor
ABVD	Austronesian Basic Vocabulary Database (Greenhill, Blust & Gray 2005-2007)	pCEP	proto-Central East Pantar
		pCP	proto-Central Pantar
AN	Austronesian	pEA	proto-East Alor
AP	Alor-Pantar	pEAM	proto-East Alor Montane
B	Blagar B > Blagar-Bama	pP	proto-Pantar
C	Central	pSWA	proto-Straits-West-Alor
		pTAP	proto-Timor-Alor-Pantar
		TAP	Timor-Alor-Pantar
		W	West
		WP	Western Pantar

## References

- Aoki, Eriko & Satoshi Nakagawa. 1993. Endenese-English dictionary. Unpublished manuscript.
- Arka, I Wayan, Fransiscus Seda, Antonius Gelang, Yohanes Nani & Ivan Ture. 2007. *A Rongga-English dictionary with an English-Rongga finderlist*. <http://chl.anu.edu.au/linguistics/projects/iwa/Web-Pages/RonggaDictionary2007.pdf>.
- Arndt, Paul. 1933. *Li'onesisch-Deutsches Wörterbuch*. Ende-Flores: Arnoldus-Druckerei.
- Arndt, Paul. 1961. *Wörterbuch der Ngadhasprache* (Studia Instituti Anthropos 15). Posieux (Fribourg, Suisse): Anthropos-Institut.
- Baird, Louise. 2002. *A grammar of Kéo: an Austronesian language of East Nusantara*. Australian National University PhD thesis.
- Baird, Louise. 2008. *A grammar of Klon: a non-Austronesian language of Alor, Indonesia*. Vol. 596 (Pacific Linguistics). Canberra: Pacific Linguistics, Research School of Pacific & Asian Studies, The Australian National University. ISBN: 9780858835948.
- Bani, Heronimus & Charles E. Grimes. 2011. Ethno-mathematics in amarasi: how to count 400 ears of corn in 60 seconds. Handout from talk given at the International Conference on Language Documentation and Conservation.

- Barnes, Robert. 1974. *Kédang: a study of the collective thought of an Eastern Indonesian people*. Oxford: Clarendon Press.
- Barnes, Robert. 1982. Number and number use in Kédang, Indonesia. *Man* (17) (1).
- Blust, Robert. nd. *Blust's Austronesian Comparative Dictionary*. <http://www.trussel2.com/acd/>.
- Blust, Robert A. 2009. *The Austronesian languages*. Vol. 602 (Pacific Linguistics). Canberra: Research School of Pacific & Asian Studies, Australian National University. ISBN: 9780858836020.
- Calon, L. F. 1890. Woordenlijstje van het dialect van Sikka. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 33. 501–530.
- Capell, Arthur. 1976. Austronesian and Papuan 'mixed' languages: general remarks. In Stephen A. Wurm (ed.), *New Guinea area languages and language study*, vol. 2 (Pacific Linguistics C40), 527–579. Canberra.
- Comrie, Bernard. 1992. Balto-Slavonic. In Jadranka Gvozdanović (ed.), *Indo-european numerals* (Trends in Linguistics 57), 717–833. Berlin & New York: Mouton de Gruyter.
- Comrie, Bernard. 2005a. Endangered numeral systems. In Jan Wohlgemuth & Tyko Dirksmeyer (eds.), *Bedrohte Vielfalt: Aspekte des Sprach(en)tods [endangered diversity: aspects of language death]*, 203–230. Berlin: Weißensee Verlag.
- Comrie, Bernard. 2005b. Numeral bases. In Martin Haspelmath, Matthew S. Dryer, Bernard Comrie & David Gil (eds.), *World atlas of language structures*, 530–533. Oxford: Oxford University Press.
- Donohue, Mark. 2008. Complexities with restricted numeral systems. *Linguistic Typology* 12. 423–429.
- Evans, Nicholas. 2009. Two plus one makes thirteen: senary numerals in the morehead-maró region. *Linguistic Typology* 13. 321–335.
- Forth, Gregory. 1981. *Rindi: an ethnographic study of a traditional domain in Eastern Sumba*. The Hague: Martinus Nijhoff.
- Forth, Gregory. 2004. *Nage birds: classification and symbolism among an Eastern Indonesian people*. London & New York: Routledge.
- Greenberg, Joseph H. 1978. Generalizations about numeral systems. In Joseph H. Greenberg (ed.), *Universals of human language*, vol. 3, 250–295. Stanford: Stanford University Press.
- Greenhill, Simon J., Robert Blust & Russell D. Gray. 2005–2007. *The Austronesian Basic Vocabulary Database (ABVD)*. <http://language.psy.auckland.ac.nz/austronesian/>.
- Grimes, Charles E., Ayub Ranoh & Helena Aplugi. 2008. *Lil Dhao (Ndao) online dictionary*. Kupang. <http://e-kamus2.org/index.html>.

- Hägerdal, Hans. 2012. *Lords of the land, lords of the sea: conflict and adaptation in early colonial Timor, 1600-1800*. Leiden: KITLV Press.
- Hammarström, Harald. 2010. Rarities in numeral systems. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rethinking universals: how rarities affect linguistic theory*, 11–60. Berlin, New York: Mouton de Gruyter.
- Hanke, Thomas. 2010. Additional rarities in numeral systems. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rethinking universals: how rarities affect linguistic theory*, 61–90. Berlin & New York: Mouton de Gruyter.
- Heine, Bernd. 1997. *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- Holton, Gary. nd. *Western Pantar lexicon*. Accessed 06 December, 2011. <http://www.uaf.edu/alor/langs/western-pantar/lexicon/>.
- Holton, Gary & Laura C. Robinson. this volume. The internal history of the Alor-Pantar language family. In Marian Klamer (ed.), *The Alor-Pantar languages*. Berlin: Language Science Press.
- 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.
- Hull, Geoffrey. 2002. *Waimaha (Waima'a)* (East Timor Language Profiles 2). Dili: Instituto Nacional de Linguística/Universidade Nacional de Timor-Leste.
- Keraf, Gregorius. 1978. *Morfologi dialek Lamalera*. Jakarta: Universitas Indonesia PhD thesis.
- Klamer, Marian. 2010. *A grammar of Teiwa*. Vol. 49 (Mouton Grammar Library). Berlin: Mouton de Gruyter.
- Klamer, Marian. 2011. *A short grammar of Alorese (Austronesian)*. Vol. 486 (Languages of the World/Materials). München: Lincom GmbH.
- Klamer, Marian. 2012. Papuan-Austronesian language contact: Alorese from an areal perspective. In Nicholas Evans & Marian Klamer (eds.), *Melanesian 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.
- Klamer, Marian, Antoinette Schapper, Greville Corbett, Gary Holton, František Kratochvíl & Laura C. Robinson. this volume. Numeral words and arithmetic operations in the Alor-Pantar languages. In Marian Klamer (ed.), *The Alor-Pantar languages*. Berlin: Language Science Press.
- Kratochvíl, František. 2007. *A grammar of Abui: a Papuan language of Alor*. Utrecht: LOT.

- Laycock, D. C. 1975. Observations on number systems and semantics. In Stephen A. Wurm (ed.), *Papuan languages and the New Guinea linguistic scene*, 219–233. Canberra: Pacific Linguistics.
- Lean, Glendon. 1992. *Counting systems of Papua New Guinea and Oceania*. Papua New Guinea University of Technology PhD thesis. <http://www.uog.ac.pg/glec/thesis/thesis.htm>.
- Lynch, John. 2009. At sixes and sevens: the development of numeral systems in vanuatu and new caledonia. In Bethwyn Evans (ed.), *Discovering history through language: papers in honour of malcolm ross*, 391–411. Canberra: Pacific Linguistics.
- Majewicz, Alfred F. 1981. Le rôle du doigt et de la main et leurs désignations dans la formation des systèmes particuliers de numération et de noms de nombres dans certaines langues. In Fanny de Sivers (ed.) (ed.), *La main et les doigts dans l'expression linguistique*, vol. 2 (LACITO - Documents Eurasie 6), 193–212. Paris: SELAF.
- Majewicz, Alfred F. 1984. Le rôle du doigt et de la main et leurs désignations en certaines langues dans la formation des systèmes particuliers de numération et des noms de nombre. *Lingua Posnaniensis* 28. 69–84.
- Matisoff, James A. 1995. Sino-Tibetan numerals and the play of prefixes. *Bulletin of the National Museum of Ethnology (Osaka) [Kokuritsu Minzokugaku Hakubutsukan Kenkyuu Hookoku]* 20(1). 105–252.
- Middelkoop, P. 1950. Proeve van een Timorese grammatica. *Bijdragen tot Taal-, Land- en Volkenkunde* (106). 375–517.
- Nishiyama, Kunio & Herman Kelen. 2007. *A grammar of Lamaholot Eastern Indonesia: the morphology and syntax of the Lewoingu dialect*. Vol. 467 (Languages of the World/Materials). München: Lincom.
- Pampus, Karl-Heinz. 2001. *Mue moten koda kiwan: kamus bahasa lamaholot, dialek lewolema, flores timur*. Frankfurt: Frobenius-Institut Frankfurt am Main.
- Pareira, M. Mandalangi & E. Douglas Lewis. 1998. *Kamus Sara Sikka - Bahasa Indonesia*. Ende, Flores, Indonesia: Penerbit Nusa Indah.
- Penn, David. 2006. *Introducing Dadu'a*. University of New England PhD thesis.
- Plank, Frans. 2009. Senary summary so far. *Linguistic Typology* 13. 337–345.
- Samely, Ursula. 1991. *Kedang (Eastern Indonesia): some aspects of its grammar*. Vol. 46 (Forum phoneticum). Hamburg: Helmut Buske. ISBN: 9783875480160.
- Sawardo, P. Wakidi, Y. Lita Tarno & S. Kusharyanto. 1987. *Struktur Bahasa Lio*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Schapper, Antoinette. 2010. *Bunaq: a Papuan language of central Timor*. Canberra: Australian National University PhD thesis.

- Schapper, Antoinette & Harald Hammarström. 2013. Innovative numerals in Austronesian languages outside of Oceania. *Oceanic Linguistics* 52(2). 425–456.
- Schapper, Antoinette & Rachel Hendery. 2014. Wersing. In Antoinette Schapper (ed.), *Papuan languages of Timor-Alor-Pantar: sketch grammars*, vol. 1, 439–504. Berlin: Mouton de Gruyter.
- 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*. Berlin: Language Science Press.
- Sidwell, Paul. 1999. The Austroasiatic numerals from ‘one’ to ‘ten’ from a historical and typological perspective. In Jadranka Gvozdanović (ed.), *Numeral types and changes worldwide*, 253–271. Berlin & New York: Mouton de Gruyter.
- Stokhof, W. A. L. 1975. *Preliminary notes on the Alor and Pantar languages (East Indonesia)*. Vol. 43 (Pacific Linguistics: Series B). Canberra: Research School of Pacific & Asian Studies, Australian National University. ISBN: 9780858831247.
- Van Klinken, Catharina. 1999. *A grammar of the Fehan dialect of Tetun, an Austronesian language of west Timor*. (C-155). Canberra: Pacific Linguistics.
- Vatter, E. 1932. *Ata Kiwan. unbekannte Bergvölker im tropischen Holland*. Leipzig: Bibliographisches Institut.
- Verheijen, Jilis A. J. 1967. *Kamus Manggarai: Manggarai-Indonesia & Indonesia-Manggarai*. Vol. 1. 's-Gravenhage: Nijhoff.
- Verheijen, Jilis A. J. 1970. *Kamus Manggarai: Manggarai-Indonesia & Indonesia-Manggarai*. Vol. 2. 's-Gravenhage: Nijhoff.
- Verheijen, Jilis A. J. 1978. *Bahasa Rembong di Flores Barat*. Vol. 3. Ruteng: S.V.D.
- Verheijen, Jilis A. J. 1982. *Komodo: het eiland, het volk en de taal*. Vol. 96 (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde). The Hague: Martinus Nijhoff.
- Wellfelt, Emilie & Antoinette Schapper. 2013. *Memories of migration and contact: East Timor origins in Alor*. Paper read at the Eighth International Convention of Asia Scholars, June 24–27, Macao.
- Winter, Werner. 1969. Analogischer Sprachwandel und semantische Struktur. *Folia Linguistica* 3. 29–45.



# Language index

Abau, 32\*

Abui, 2, 5, 9, 14, 17–19, 21–23, 25, 27, 36, 37, 50, 51, 55–62, 64–67, 70–73, 77, 78, 80–87, 122, 124, 125, 127, 129–138, 152, 157, 169, 172, 174, 214–217, 224–228, 230, 231, 232\*, 235–240, 249, 250\*, 250, 251, 263–266, 267\*, 267, 270, 271, 277, 278, 289\*, 290, 292, 295, 304, 306, 308, 320, 321, 323–325, 338, 339, 340\*, 340, 343, 344, 347–350, 353–355, 357, 358, 361, 363, 374, 375, 379, 389\*, 389–392, 395–398, 400–408, 411–414, 416–419, 423–430, 436, 442, 443

Adang, 5, 9, 14, 20, 26, 27, 32, 36, 50\*, 51, 55–57, 60, 61, 62\*, 62–67, 70–73, 76, 77, 79–87, 104, 122–125, 127, 129–138, 162, 163, 166, 169, 172, 174, 178, 211–214, 224–228, 230, 231, 235–238, 240, 245, 254–258, 263–266, 267\*, 267–271, 276–278, 290, 292, 298, 304, 306, 308, 319–321, 323–326, 338, 339, 342, 344, 346, 348, 353, 354, 356, 358, 360, 362, 375, 377, 378

Alorese, 2, 4, 7, 9, 11, 50, 63, 79, 86,

236, 238–240, 291, 314, 315, 321, 327, 328

Amarasi, 321, 329, 330

Angave, 33\*

Apui, 8

Atauro, 307, 321, 329, 330

Austronesian language(s), 1, 2, 4, 6, 7, 10, 11, 13, 31, 33, 34, 50\*, 50, 58, 61–63, 70, 86, 87, 94, 135, 138, 143, 150, 164, 174, 177, 180, 233, 236, 238–240, 284, 286, 289\*, 289, 291, 294, 300, 301, 303, 306, 307, 309–316, 321, 322, 327, 329

Awará, 33\*

Baham, 180, 182, 183, 185

Bajau, 7

Barai, 160

Bird's Head language(s), 150, 152

Blagar, 5, 9, 14, 15, 20, 23, 27, 36, 51, 55–57, 60–65, 67, 70–73, 76, 77, 79–87, 104, 122–125, 127, 129–138, 169, 172–174, 181, 196, 197, 205, 207\*, 207–209, 224–229, 231, 232\*, 233, 234\*, 236, 239, 240, 251\*, 251–253, 258, 263–265, 267\*, 267–270, 276–278, 289\*, 290, 292, 298, 304, 306, 308, 319–321, 323–326, 331, 338, 342, 358

- Bomberai language(s), 150, 153, 175,  
180, 182, 185–187
- Bunaq, 6, 31, 93–114, 116–119, 123, 125,  
127, 139–143, 156, 157, 162,  
168, 169, 172, 173, 300, 321
- Chambri, 33\*
- Chenapian, 33\*
- Dadu'a, 321, 329, 330
- Dani language(s), 150, 153, 175, 178–  
180
- Deing, 5, 14, 36, 78, 127, 129–138, 290,  
292, 295, 304, 308, 320, 321,  
323–326, 393
- Dhao, 321, 329, 330
- Dutch, 285
- Dyirbal, 243
- East Caucasian language(s), 243
- Ekari, 179, 180
- Ende, 312, 314, 315, 321, 327, 328
- Enga, 160
- Fataluku, 6, 93, 95–109, 111–114, 116–  
119, 127, 139–143, 169, 172, 173,  
174\*, 174
- Finisterre-Huon language(s), 158
- Folopa, 33\*
- Fore, 158, 160
- Galela, 169
- Guarjío, 243
- Hamap, 5, 36, 127, 129–138, 174, 178,  
290, 292, 298, 304, 308, 319–  
321, 323–326
- Hamtai, 52
- Hindi, 413
- Idate, 321
- Iha, 180–182, 184, 185
- Ilongot, 285, 286, 321
- Inanwatan, 160, 161, 176, 180
- Indonesian, 2, 8, 9, 13, 30, 33, 55\*, 79,  
337, 350, 354, 357, 358, 369
- Iwam, 32\*
- Kéo, 312, 314, 321, 327, 328
- Kabola, 5, 9, 36, 62\*, 127, 129–138, 173,  
174, 178, 238, 290, 292, 298,  
304, 308, 319–321, 323–326
- Kaera, 5, 8, 9, 15, 16, 20, 23, 26, 28, 36,  
37, 51, 52, 54–57, 59–67, 70–  
72, 76, 77, 80–87, 122, 124,  
125, 127, 129–138, 169, 172,  
180, 182, 266, 267\*, 267–270,  
277, 278, 290, 292, 295, 296,  
304, 306, 308, 320, 321, 323–  
326, 340\*, 342, 358, 360, 374
- Kafoa, 5, 9, 36, 127, 129–138
- Kamang, 2, 5, 9, 14, 19, 21, 23, 25, 27,  
28, 32, 36, 37, 51, 55–64, 66,  
67, 70, 72, 73, 76–78, 80–87,  
122–125, 127, 129–138, 172, 174,  
196, 216, 218, 219\*, 219, 220,  
224–228, 230, 231, 232\*, 233,  
235–240, 245, 261–265, 267\*,  
267–272, 274, 276–278, 287,  
288, 289\*, 290–293, 295, 303–  
306, 308, 320, 321, 323–326,  
338, 340\*, 341, 343, 344, 347–  
350, 353–355, 357, 358, 361,  
363, 368, 374, 375, 379, 385–  
387, 388\*, 388, 389, 393, 394\*,  
394–397, 400–403, 405–407,  
411–416, 418, 419, 422, 427,  
429, 431–436, 442, 443
- Kambera, 86, 300
- Karas, 180, 185, 186



- Kedang, 233, 291, 300, 312, 314, 321, 327, 328
- Kemak, 307, 312, 321, 329, 330
- Keo, 315
- Kiraman, 5, 208, 210, 211, 224–228, 230, 231, 233, 235, 237, 240
- Klon, 5, 9, 14, 17, 19, 20, 22, 23, 27, 32, 36, 37, 51, 55–57, 60, 62\*, 62–67, 70–73, 76, 77, 80–86, 122, 124, 125, 127, 129–138, 169, 172, 174, 178, 180, 232\*, 266, 267\*, 268–270, 277, 288, 290, 292, 298, 302, 304–306, 308, 319–321, 323–326, 374–376, 380\*, 408
- Kokoda, 176, 177
- Komodo, 321, 327, 328
- Kui, 5, 12, 36, 51, 55–57, 59, 60, 62–67, 70, 72, 73, 76, 77, 80–87, 122, 127, 129–138, 152, 169, 172, 174, 178, 180, 266, 267\*, 268, 270, 277, 278, 290, 292, 297, 298, 301, 303, 304, 307, 308, 315, 317, 319–321, 323–326
- Kula, 5, 9, 36, 124, 127, 129–138, 272\*, 290–293, 295, 304–308, 312, 320, 322–326, 339, 341, 343, 358, 374, 375, 380\*
- Kunimaipa, 52
- Lakalei, 322, 329, 330
- Lamaholot, 7, 86, 307, 314, 315, 322, 327, 328
- Lio, 312, 314, 315, 322, 327, 328
- Lower Grand Valley Dani, 178
- Makalero, 6, 31, 93, 95, 97–105, 107–109, 111–114, 116–119, 124, 125, 127, 139–143, 168, 169, 172, 174\*, 175, 277
- Makasae, 6, 93, 95, 97–105, 107–109, 111–114, 116–119, 127, 139–143, 169, 172, 174\*, 175, 232\*, 277
- Malay, 2, 8, 9, 13, 33, 61, 62, 135, 239, 289\*, 289, 291, 307, 354, 357, 364–368
- Mambai, 86, 174, 294, 300, 307, 312, 314, 322, 329, 330
- Manggarai, 174, 322, 327, 328
- Marind, 159, 163
- Mian, 160, 162
- Miskitu, 407
- Nage, 294, 312, 314, 322, 327, 328
- Naueti, 312
- Nedebang, 5, 36, 51, 52, 55–61, 63–68, 70–72, 76, 77, 80–87, 122, 127, 129–138, 169, 182, 224
- New Guinea language(s), 243
- Ngadha, 312, 314, 315, 322, 327, 328
- North Halmahera language(s), 163
- North Halmaheran language(s), 150, 152–154, 156, 158, 159, 161–164, 166–169, 171, 186
- Oirata, 6, 50, 93, 95, 97–105, 107–109, 111–114, 116–119, 127, 139–143, 169, 172, 173, 174\*, 175
- Old Malay, 58
- Palu'e, 322, 327, 328
- Portuguese, 13
- pronouns, 180
- proto-Alor, 169, 174, 178
- proto-Alor-Pantar, 15, 17, 22, 23, 28, 30–32, 50–61, 64–71, 73, 77, 79–88, 94–109, 110\*, 110–124,

- 128–138, 144–146, 165, 166,  
168, 169, 171\*, 172–174, 176,  
177, 179–183, 185, 186, 224,  
228, 231, 236, 239, 240, 244,  
266, 267\*, 267–272, 274, 276,  
278, 284, 287, 288, 289\*, 289–  
293, 296\*, 297, 299–301, 303,  
305, 306, 310, 316–320, 340\*,  
342, 374, 379, 390, 407, 408  
proto-Austronesian, 58, 62, 86, 196,  
291, 299, 303, 307\*, 307, 311–  
313, 315, 316, 327–330  
proto-Bomberai, 187  
proto-Central East Alor, 303, 317\*, 318,  
319\*, 320  
proto-Central East Pantar, 296  
proto-Central Pantar, 296, 320  
proto-East Alor, 271\*, 317\*, 318, 319\*,  
320  
proto-East Alor Montane, 317\*, 318,  
319\*, 320  
proto-Eastern Timor, 94, 97, 139  
proto-Malay, 316  
proto-Malayo Polynesian, 173  
proto-Malayo-Polynesian, 33, 62, 86,  
87, 135, 138, 143  
proto-North-Halmahera, 155, 164–166,  
168, 170  
proto-Pantar, 296, 297, 317\*, 318, 320  
proto-Straits-West-Alor, 297–303, 316,  
317\*, 318, 319  
proto-Straits-West-Pantar, 320  
proto-Timor, 30, 94–109, 110\*, 110–  
120, 122, 128, 139–146, 172,  
174, 178\*, 180, 182  
proto-Timor Alor Pantar, 30, 52, 93–  
100, 102–107, 109–116, 118, 120–  
122, 128, 144–146, 150\*, 155,  
167–170, 171\*, 171, 173, 174,  
177–182, 185, 186, 310  
proto-Trans-New-Guinea, 152, 155\*,  
155, 158, 159, 164, 165, 171–  
175, 179, 180, 182, 185  
proto-West Malayo-Polynesian, 289\*  
proto-Western Pantar, 302, 303  
  
Rembong, 322, 327, 328  
Reta, 5, 127, 129–138  
Retta, 173, 175, 289, 290, 292, 298, 304,  
308, 319, 320, 322–326  
Rongga, 312, 314, 315, 322, 327, 328  
Russian, 289\*  
  
Sama-Bajo, 50\*  
Sar, 5, 36, 61, 124, 127, 129–138, 290,  
292, 295, 296, 304, 308, 320,  
322–326  
Sawila, 5, 9, 13, 27, 28, 36, 37\*, 37,  
51, 55–64, 66, 67, 70, 72, 73,  
76–78, 80–87, 122–125, 127,  
129–138, 169, 173, 175, 266,  
267\*, 268–270, 272\*, 272, 273,  
276–278, 288, 290–293, 295,  
304, 306–308, 312, 320, 323–  
326, 341, 343, 358, 362, 374,  
375, 380\*, 403  
Sika, 322, 327, 328  
South Bird's Head language(s), 153,  
175, 177  
  
Tabaru, 169  
Tambora, 150\*  
Tanae, 33\*  
Tehit, 52  
Teiwa, 5, 8, 9, 14, 16–18, 20, 22–27,  
32, 36, 37, 50\*, 51, 52, 54–  
67, 70–72, 76, 77, 78\*, 80–87,

- 122–125, 127, 129–138, 152,  
156, 169, 172, 180–182, 196,  
197, 204–206, 207\*, 207, 222,  
224–229, 231, 232\*, 233, 236,  
239, 240, 248, 249, 263–265,  
267\*, 268–270, 277, 278, 290,  
292, 295, 296\*, 296, 304, 306,  
308, 309, 320, 322–326, 338–  
340, 342, 344, 346–348, 352–  
354, 356, 358, 359, 362, 374–  
379, 381–384, 385\*, 385, 390,  
392, 393, 395–402, 404–407,  
411–415, 417–419, 427, 436–  
444
- Tetun, 58, 307
- Tetun Fehan, 291, 300, 322, 329, 330
- Thao, 289\*
- Tibeto-Burman language(s), 243
- Tobelo, 158, 159, 161–163, 166
- Tokodede, 58, 294, 306, 307, 312, 314,  
315, 322, 329, 330
- Trans-New Guinea language(s), 10, 34,  
52, 150, 152–154, 156, 158–  
164, 166, 167, 171, 186
- Uab Meto, 322, 329, 330
- Ujir, 285
- Uto-Aztec language(s), 243
- Waima'a, 322, 329, 330
- Wantoat, 33\*
- Wersing, 5, 36, 37, 51, 55–58, 60–64,  
66, 67, 70, 72, 73, 76–78, 80–  
87, 122, 124, 125, 127, 129–  
138, 152, 169, 221–228, 230,  
231, 233, 236, 237, 240, 245,  
247–249, 263–265, 267\*, 268,  
269\*, 269, 270, 277, 278, 290–  
293, 295, 304, 306–309, 312,  
320, 322–326, 374–377, 379,  
392, 393, 394\*, 394–403, 405,  
407
- West Papuan language(s), 152, 153, 175
- Western Dani, 178
- Western Pantar, 5, 8, 9, 12\*, 14, 15, 20,  
21, 23, 24, 28, 32, 36, 51–53,  
55–67, 70–72, 76–78, 80–87,  
121–125, 127–138, 158–160, 163,  
169, 175, 180, 196–205, 224–  
229, 231, 233, 234, 236, 237,  
239, 240, 245, 247, 248, 258–  
261, 263–266, 267\*, 267–271,  
276–278, 289–291, 293, 294,  
297, 298, 301–304, 306–308,  
315, 317, 319, 320, 322–325,  
339, 340\*, 342, 344, 347, 348,  
352–354, 356, 358, 359, 362,  
374–377, 379, 380\*, 380, 381,  
383, 392, 395–397, 400, 401\*,  
401, 402, 404, 405, 407, 408,  
418
- Wissel Lakes language(s), 150, 153, 175,  
179
- Wogamusin, 33\*

# Subject index

- actor, 417, 418  
additive numeral, 284, 287, 294, 295,  
301, 303, 305\*, 305–307, 311,  
315–317  
adposition, 16, 26, 29, 162, 164, 249,  
251, 305  
affectedness, 411–414, 416, 417, 422,  
425–427, 435, 442, 443  
agreement, 442  
alienability, 22, 23, 29, 60, 102, 106,  
122, 128, 160, 162, 199, 342,  
352  
alignment, 19, 20, 25, 35, 156, 414  
animacy, 18, 338, 376, 399, 405, 406,  
412–415, 417, 420, 423, 427,  
429, 435–439, 441–444  
apocope, 63, 65, 95, 98  
aspect, 24, 25, 28, 162  
  
borrowing, 13, 14, 31, 33, 50\*, 58–61,  
63–66, 67, 68\*, 70, 76, 78, 86,  
87, 94, 121, 135, 138, 143, 166,  
168, 177, 178\*, 180, 186, 236,  
284, 289\*, 289, 291, 294, 299–  
301, 303, 306, 307, 312, 314–  
317, 319, 337, 340, 364  
  
cardinal numeral(s), 284, 285, 293, 297,  
305, 309, 314, 316, 319, 323,  
338–341, 344–347, 348\*, 348–  
350, 353, 354, 357, 364, 366–  
369, 371  
  
clause-chaining, 29, 30, 163  
conjunction, 16, 24, 29, 408  
conventionalization, 443, 444  
  
definiteness, 421  
deixis, 35, 97, 128, 243–247, 251\*, 258,  
260, 261, 271, 276, 278  
demonstrative, 245, 246\*, 247, 249,  
250\*, 250, 251\*, 251–255, 257–  
260, 263, 269, 278, 379, 381–  
383, 385, 390, 408  
derivation, 7  
determiner, 16  
dialect chain, 153  
differential object marking, 413  
differential subject marking, 413  
direction, 24, 245, 254–257, 261, 263,  
264, 269, 271, 274, 276  
distributive, 164  
distributive numerals, 338, 345–347,  
348\*, 348–350, 353, 357, 359,  
362, 363  
  
elevation, 97, 243, 244\*, 244–248, 249\*,  
249, 250\*, 250–252, 254–268,  
269\*, 269–271, 272\*, 272–277  
epenthesis, 59, 63–65, 121  
  
free pronoun, 412, 414, 416, 427, 428  
gender, 30  
inflection, 7

- inflection classes, 423, 431, 434
- innovation, 31, 32, 50, 58, 61, 63, 67, 70, 71, 73, 74, 76\*, 76, 78, 79, 94, 97, 153, 239, 272, 274, 276, 293, 297, 302, 311, 312, 314, 317, 319, 320, 374, 379, 407
- isolate language(s), 150, 153, 186
- kinship, 34, 35, 101, 195, 197, 199, 202, 205, 206, 207\*, 208, 210, 211, 213, 217, 218, 220, 223–227, 231, 233, 236, 238–240, 404
- lexical stipulation, 412, 423, 436, 443
- manner, 24
- metathesis, 64, 86, 87, 95, 98, 100, 103, 105, 115, 129, 139–141, 153, 180
- metatypy, 78
- mood, 28
- motion, 244, 245, 247–252, 254, 255, 258, 260–264, 266, 267, 271, 272\*, 272–276
- multiplicative numeral, 284, 287, 301, 302, 311, 315
- negation, 16, 29, 163
- nominalisation, 97, 385
- numeral classifier, 31–33, 312, 313, 327, 328, 338, 369, 370, 378, 379, 381, 383, 384, 386, 408
- Nusa Tenggara Timor, 8
- object, 415, 417, 436–439, 441
- ordinal numeral(s), 305, 307, 338, 341–345, 357, 359–362, 364, 365, 368, 369
- palatalisation, 62\*, 62
- person, 414, 418\*, 418, 441
- plural (number) word, 373–376, 377\*, 377–383, 385\*, 386–393, 394\*, 394–408
- plurality, 23, 35, 60, 102, 128, 164, 176, 178, 186, 370
- possession, 7, 22, 23, 26, 28, 29, 60, 102, 106, 122, 128, 156, 159, 160, 162, 183, 198, 199, 201, 212, 215, 278, 338, 341–343, 345, 357, 370, 378, 380, 396, 404, 408
- prefix alternation, 424, 425, 432, 438, 441
- productivity, 437
- pronominal indexing, 411–414, 425, 443
- pronoun, 7, 8, 16, 29, 30, 150, 152, 156–158, 160, 162, 164–167, 176, 178, 179, 181, 183, 186, 198, 201, 212, 340, 341, 345, 370, 376\*, 379, 382, 383, 385\*, 387, 388\*, 388, 389, 391, 393, 394\*, 394, 396, 398, 402, 405–407
- reduplication, 103, 294, 302, 338, 345–350, 357, 368, 408
- referential properties, 411, 443
- right-headed, 15, 161
- secundative alignment, 442\*
- semantic alignment, 412, 414–416, 420, 427–429, 443
- serial verb construction(s), 24, 25, 27, 29, 350
- simplex numeral, 285, 291, 293, 294, 310\*, 310, 316, 317
- sound change, 284, 287, 288\*, 288, 319
- specificity, 413, 421
- subject, 417, 443

## *Subject Index*

- subtractive numeral, 284, 287, 294,  
297, 299–303, 310, 311, 316,  
317
- switch reference, 29, 30
- syntactic alignment, 412, 414, 436
- telicity, 420
- tense, 28
- undergoer, 417, 418
- valency, 27
- verb classes, 412, 423, 429, 436, 438,  
439, 444
- verb-final, 15, 161
- video stimuli, 420
- volitionality, 157, 411–414, 416, 417, 420,  
422, 427–429, 442, 443
- vowel harmony, 87



# The Alor-Pantar languages

The Alor-Pantar family constitutes the westernmost outlier group of Papuan (Non-Austronesian) languages. Its twenty or so languages are spoken on the islands of Alor and Pantar, located just north of Timor, in eastern Indonesia. Together with the Papuan languages of Timor, they make up the Timor-Alor-Pantar family. The languages average 5,000 speakers and are under pressure from the local Malay variety as well as the national language, Indonesian. Together with the Papuan languages of Timor, they make up the Timor-Alor-Pantar family.

This volume studies the internal and external linguistic history of this interesting group, and showcases some of its unique typological features, such as the preference to index the transitive patient-like argument on the verb but not the agent-like one; the extreme variety in morphological alignment patterns; the use of plural number words; the existence of quinary numeral systems; the elaborate spatial deictic systems involving an elevation component; and the great variation exhibited in their kinship systems.

Unlike many other Papuan languages, Alor-Pantar languages do not exhibit clause-chaining, do not have switch reference systems, never suffix subject indexes to verbs, do not mark gender, but do encode clusivity in their pronominal systems. Indeed, apart from a broadly similar head-final syntactic profile, there is little else that the Alor-Pantar languages share with Papuan languages spoken in other regions. While all of them show some traces of contact with Austronesian languages, in general, borrowing from Austronesian has not been intense, and contact with Malay and Indonesian is a relatively recent phenomenon in most of the Alor-Pantar region.