

7 *Numeral classifiers in the Papuan languages of Alor and Pantar*

A comparative perspective

MARIAN KLAMER¹

1 Introduction

When a language forms a new class of numeral classifiers, this is often due to a language internal process of grammaticalization (cf. Aikhenvald 2000). In this paper I argue that this process of grammaticalization can be both internally motivated and modelled after functions that speakers encounter through contact with other languages.²

On the basis of recently collected survey data I investigate the history of the numeral classifiers in six Papuan languages, spoken on the islands of Alor and Pantar. The six languages of my sample cover most of the Alor-Pantar region, and represent the major subgroups of the Alor-Pantar family. I argue (i) that the classifiers found in the languages today have not been inherited from proto-Alor-Pantar, but are a later development in each of the languages; and (ii) that this development must have taken place under influence from Austronesian languages with classifiers.

Numeral classifiers are “morphemes that only appear next to a numeral, or a quantifier; they may categorize the referent of a noun in terms of its animacy, shape, and other inherent properties” (Aikhenvald 2006:466). Two basic types of numeral classifiers are generally distinguished: mensural and sortal classifiers. A mensural classifier “individuates in terms of quantity”, and a sortal classifier “individuates whatever it refers to in terms of the kind of entity that it is” (Lyons 1977:463). Most, if not all, languages have mensural classifiers, while the worldwide distribution of sortal classifiers is more restricted (see Gil 2011). In this paper, the term ‘classifier’ will refer to sortal numeral classifiers only; mensural classifiers will not feature here.

In the Alor-Pantar archipelago, some 20 Papuan (‘non-Austronesian’) languages are spoken (see Map 1). Together these form the Alor-Pantar (AP) subgroup of the Timor-Alor-Pantar (TAP) family (Holton et al. 2012, Schapper et al. 2012). There is evidence that the Alor-Pantar languages originate from the Straits region between the two islands (Robinson and Holton 2012). The Papuan character of the TAP languages has long been recognized in the literature. Until recently, most authors have assumed that the TAP family belongs to the putative Trans-New Guinea family, but as there is no lexical evidence

¹ ✉ M.A.F.Klamer@hum.leidenuniv.nl

² I wish to thank members of the audience at the ICAL conference in Bali in July 2012 for their valuable feedback on the oral presentation of this paper, and the two anonymous reviewers for helpful comments on the written version.

supporting this position, Holton et al. (2012) instead propose that the TAP group should be considered a distinct family, unrelated to Trans-New Guinea, which is also the position taken here. The TAP family appears to be relatively young; calculations by Holman et al. (2011) suggest it to be some 3,500 years old. There is good evidence that the Alor-Pantar languages have been in contact with Austronesian languages since prehistoric times: Austronesian loans have been reconstructed back to proto-Alor-Pantar (Holton et al. 2012: 114) and there is Austronesian influence in Alor-Pantar numeral systems (Schapper and Klamer 2014).

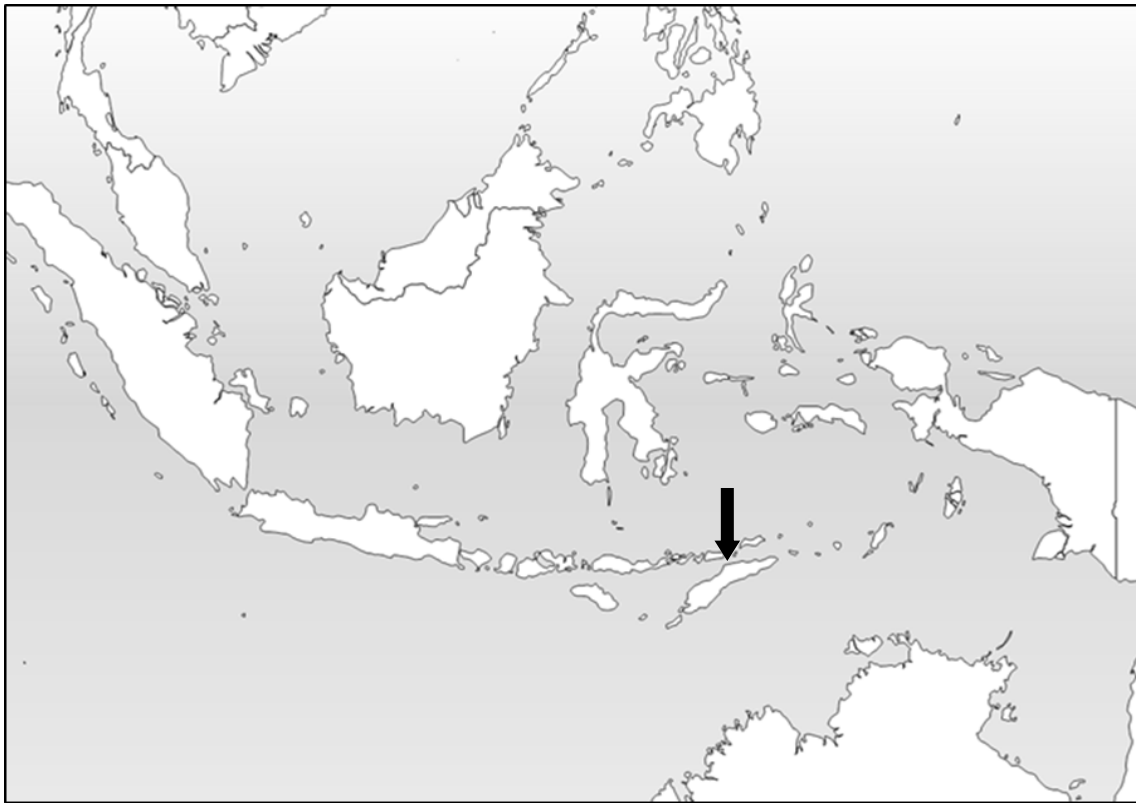


Figure 1. Location of the Alor-Pantar languages in Eastern Indonesia

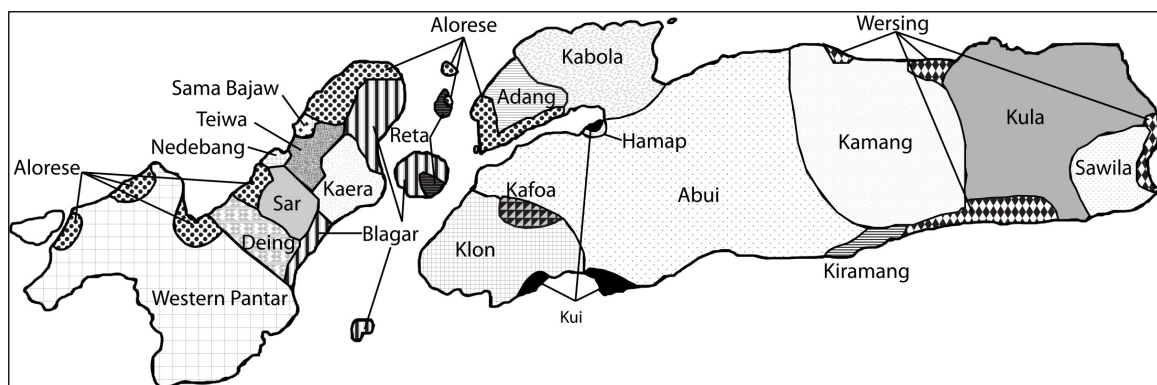


Figure 2. Map of the Alor-Pantar languages

This paper is structured as follows. In section 2 I present the classifiers in my sample of six Alor Pantar languages, and argue that there is no evidence to reconstruct one or more classifiers for proto-Alor-Pantar. Instead, they must be the result of different types of relatively autonomous developments in individual languages that yet need to be scrutinized. In section 3, I briefly review the literature suggesting that numeral classifiers

are not a feature typically found in Papuan languages. However, the Papuan languages that do have numeral classifiers are found in areas where Austronesian-Papuan contact existed. This indicates that classifiers in Papuan languages may be due to contact (section 3.1). I then present some evidence that classifiers are typically found in Austronesian, which supports the contact scenario (section 3.2). In section 4 I summarize the conclusions.

2 Numeral classifiers in Alor Pantar are not inherited

Numeral classifiers are found throughout the Alor-Pantar family, in languages spoken across the two islands. In this section, I argue that no classifier is reconstructable for proto-Alor Pantar, and that classifiers in Alor Pantar probably developed out of nouns.

Illustrations of numeral classifiers in the sample languages are given in Table 1-6 below. The column ‘Meaning’ represents the lexical meaning of source of the numeral classifier; in this column, the value ‘none’ means that today’s speakers do not attribute a meaning to the form, while ‘not reported’ indicates that the source does not include information about the meaning.

Tables 1-6 present the languages in geographical west-to-east order: Teiwa (Table 1) and Western Pantar (Table 2) on Pantar island; Adang (Table 3) and Klon (Table 4) on the western part of Alor; Abui in central Alor (Table 5) and Kamang in central-east Alor (Table 6), see Figure 1 for language locations.

Table 1. Classifiers in Teiwa (Klamer 2010, 2014)

Form	Meaning	Classifies
<i>bag</i>	‘seed’	general classifier for everything but humans and fruits/tubers
<i>-man3</i>	(none)	humans
<i>kam</i>	‘long fruit’	long fruits, e.g. tamarind, banana
<i>yis</i>	‘cylindrical fruit’	cylindrical fruits and tubers, e.g. taro, cassava
<i>quu</i>	‘round fruit’	round fruits, e.g. mango, papaya, lemon, pumpkin, coconut

Table 2. Classifiers in Western Pantar (Holton, this volume; to appear)

Form	Meaning	Classifies
<i>bina</i>	‘detached’ ⁴	general classifier
<i>haila</i>	‘base, area’	objects with areal extent
<i>dawal</i>	‘roll, coil’	rope-like objects (e.g. rope, nylon, cable thread)
<i>dis</i>	‘stringer’	objects strung on a string
<i>kakka</i>	‘stiff’	long, stiff, flat objects
<i>gamma</i>	‘nose, point of land’	sticklike objects
<i>hissa</i>	‘fruit’	fruit, contents
<i>kassi</i>	‘to split’	split-off objects
<i>lu’a</i>	‘rounded, oblong’	rounded object
<i>waya</i>	‘leaf’	flat, flexible objects

³ This form must take an obligatory person-marking prefix.

⁴ Western Pantar is unique among the languages surveyed here in recruiting classifiers from nominal as well as non-nominal lexemes.

Table 3. Classifiers in Adang (Robinson and Haan, to appear)

Form	Meaning	Classifies
<i>pa'</i>	non-round fruit	general classifier for objects of many shapes and sizes; e.g. arrows, drums, borrowed nouns, birds, fish ⁵
<i>beh</i>	'leaf'	flat, flexible objects (leafs, money notes)
<i>bo'</i>	(none)	flat, rigid, large objects, incl. fields
<i>bo'</i>	'log'	long, cylindrical, rigid objects, e.g. bamboo, logs
<i>'ahang</i>	'slice'	flat, rigid, small objects, e.g. wood, walls
<i>'anemeng</i>	'sheet'	non-flat, flexible objects, e.g. clothes, rope, string
<i>el</i>	(none)	rigid, standing objects, e.g. buildings, trees
<i>'afail</i>	'seed'	small rigid objects, e.g. corn kernel, rice grain
<i>'ir</i>	(none)	long, cylindrical, jointed objects, e.g. bamboo, sugarcane
<i>kumang</i>	'piece'	short, cylindrical objects, e.g. logs, eels, snakes
<i>pir</i>	'round fruit'	round objects, e.g. fruits, animals, people
<i>puh</i>	(none)	hanging objects, e.g. banana blossoms, corn ears
<i>bar</i>	'bunch'	short, clustered, hanging objects, e.g. coconut bunches, earrings, bells, betel nut
<i>buil/buling</i>	'stem'	long, clustered hanging objects, e.g. banana bunches, rice

Table 4. Classifiers in Klon (Baird 2008)

Form	Meaning	Classifies
<i>ip</i>	'amount'	objects (formal usage)
<i>up</i>	'amount'	objects (informal usage)
<i>ana</i>	(not reported)	human classifier

Table 5. Classifiers in Abui (Schapper 2010)

Form	Meaning	Classifies
<i>upi</i>	'fruit'	fruits, animals
<i>lohu</i>	'long [thing]'	bronze drums, larger animals
<i>kasing</i>	'bit'	man-made items
<i>-ning6</i>	(not reported)	human classifier

Table 6. Classifiers in Kamang (Schapper, to appear)

Form	Meaning	Classifies
<i>uh</i>	(none)	general classifier
<i>ning</i>	(none)	human classifier

In numeral NPs in the Alor-Pantar languages the classifier follows the noun, and precedes the numeral: [NOUN - CLASSIFIER - NUMERAL]. An illustration is (1), with the Teiwa general classifier *bag*:

⁵ The Adang default classifier *pa'* (*a*) is derived from a noun originally meaning 'small non-round fruit' and is now used 'to count any non-liquid object', including borrowed items, birds and fish (Haan 2001: 296).

⁶ This form must take an obligatory person-marking prefix.

Teiwa

- (1) *Qarbau bag ut ga'an u*
 water.buffalo CLF four DEM DIST
 'Those four water buffalos'

The following general observations can be made relating to the data presented in Table 1-6. First, the inventories of classifiers differ significantly in size. For instance, Adang has 14 classifiers while Kamang has only 2.

Second, the classifier lexemes differ significantly in their shape, the lexical source from which they derive, as well as their classifying function. While the set contains some parallel forms (e.g., the classifiers originating from 'leaf' in W Pantar *waya* and Adang *beh*, and the human classifier *ning* in Abui and Kamang), no cognates are found across the sample.

Third, a number of the languages have a 'general' classifier, which functions to classify nouns outside the semantic domains of the other, semantically more specific, classifiers (c.f. Zubin and Shimojo 1993, Shimojo 1997). Although these lexemes share a common general classifying function, they derive from different lexical sources: Teiwa *bag* < 'seed', Western Pantar *bina* < 'detached', Adang *pa'* < 'non-round fruit', and Kamang *uh* (meaning of source unknown), so that no proto-form for a general classifier can be constructed.

Finally, each of the languages uses its classifiers to make classifications of quite different nature. For instance, in Teiwa, fruits are not classified together, but in different classes according to the shape of the fruit (*kam*, *yis*, *quu'*), while in Adang, fruits are classified together with animals and people (*pir*), Western Pantar classifies fruits with 'contents' (*hissa*), and Klon and Kamang do not classify fruits at all. The few languages that classify animals put them in a class with fruits and humans (Adang *pir*), or with inanimate items (Abui *lohu*).

In sum, the survey data presented in this section show that (i) numeral classifiers are very common in Alor-Pantar languages, but (ii) they have variable shapes and origins, and make very different semantic classifications, so that (iii) no classifier is reconstructable for proto-Alor Pantar.

If this is correct, then the classifiers attested in the individual languages must have developed after the split up of proto-Alor Pantar. They are not inherited from the proto language, but have developed independently in the individual languages. The data presented here suggest that classifiers in Alor Pantar developed out of nouns.

It seems that nouns indicating the parts of a plant, such as 'fruit', 'leaf', and 'seed' are among the preferred sources for numeral classifiers. Illustrations of such part-of-whole nouns in Teiwa are given in (2a-d), where the generic plant name *wou* 'mango-hood'⁷ combines with the part-of-whole nouns *bag* 'seed', *wa'* 'leaf', *qaau* 'flower', and *heer* 'stem, base' in order to refer to certain particular parts of a mango-plant:

Teiwa

- (2) a. *wou bag*
 mango-hood seed
 'mango seed(s)'
- b. *wou wa'*
 mango-hood leaf
 'mango leaf(s)'

⁷ *Wou* is glossed as 'mango-hood' because (unlike what the gloss 'mango' would suggest) *wou* by itself has no referential meaning: it must combine with a fruit classifier *quu'* to refer to (a) mango fruit(s).

c. *wou* *qaau*
 mango-hood flower
 ‘mango flower(s)’

d. *wou* *heer*
 mango-hood stem
 ‘mango tree(s)’

It is plausible that classifiers such as Teiwa *bag* developed out of the part-of-whole noun ‘seed’ through structural reanalysis of an ambiguous phrase structure. This is illustrated in (3)–(5). Part-of-whole nouns like Teiwa *bag* in (5) can develop into a classifier through reanalysis of their structural position. As a part-of-whole noun, it is part of a nominal compound, (3), but such a structure may be reanalyzed into the one in (5), where *bag* is clearly functioning as a general classifier, via ambiguous structures like the one in (4). (4) is ambiguous because *bag* can be analyzed here as either a part-of-whole noun within a nominal compound *wou bag* that is followed by a NUMeral (as in (3)); or as a numeral classifier that combines with a numeral into a NUMeral Phrase *bag yerig* (as in (5)).

Teiwa

(3) [N_{HEAD} NUM]_{NP}
 / \ |
 [*wou* *bag*] _N *yerig*

(4) *wou* *bag* *yerig*
 mango-hood seed three
 ‘three mango seeds’

(5) [N_{HEAD} NUMP]_{NP}
 | / \ |
 bala’ [*bag* *yerig*]
 bed CLF three
 ‘three beds’

What these few examples are meant to illustrate here is how a numeral classifier can develop out of a part-of-whole noun through a simple (‘re-bracketing’) reanalysis of the structure of numeral NPs. I have argued elsewhere (Klamer 2014) that this is what may have happened in Teiwa. The historical trajectories of the classifiers in the other languages mentioned above still need to be investigated. However, what is clear from the data presented in this section is that nouns, especially part-of-whole nouns are among the preferred sources for the numeral classifiers that developed in the Alor-Pantar languages.

3 Numeral classifiers in Alor Pantar as an Austronesian feature

3.1 Numeral classifiers in Papuan languages

From a Papuan perspective, it is interesting to find that classifiers have developed across the Alor-Pantar family. Numeral classifiers are not at all typical for Papuan languages. Neither Foley (1986, 2000) nor Aikhenvald and Stebbins (2007) mention numeral classifiers among the features that are typical for Papuan languages. In addition, the numeral classifiers map by Gil (2011) lists two dozens of Papuan languages across New Guinea, and all of them lack numeral classifiers. This does not mean that there are no

Papuan languages with classifiers at all: Aikhenvald (2000:123) mentions ten such languages.⁸ However, these languages occur in scattered locations, and are members of different Papuan families in the eastern part of New Guinea. In contrast to this, there are geographical clusters of Papuan languages with numeral classifiers on the western side of New Guinea, as shown in Figure 3. Apart from the languages of Alor and Pantar, we find Papuan languages with classifiers on Timor, in the Bird's Head of mainland New Guinea, as well as in Halmahera.

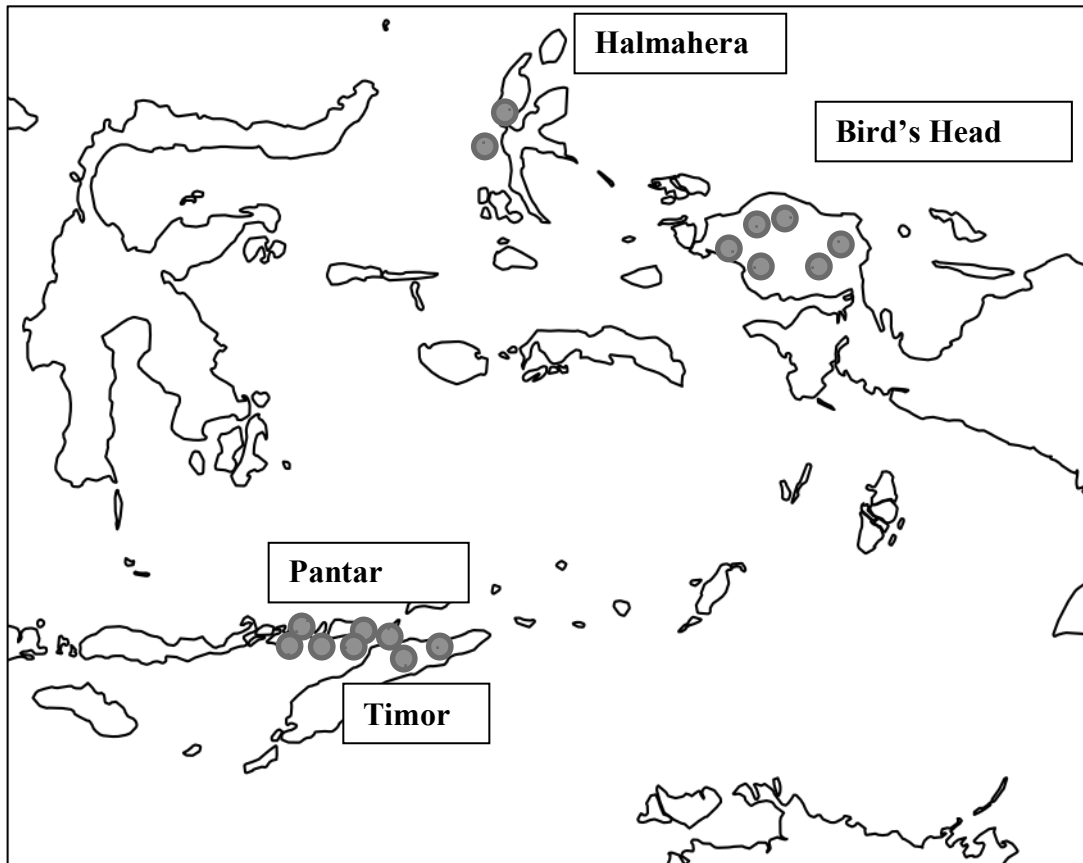


Figure 3. Papuan languages with numeral classifiers in eastern Indonesia, represented as grey dots

These areas are exactly the ones where we know that longstanding and intense contacts between Papuan and Austronesian speaking groups have existed, and resulted in diffusion of Austronesian features into Papuan languages (e.g. the inclusive-exclusive person distinction), or Papuan features into Austronesian (e.g. the alienable-inalienable possessor distinction).⁹

On Timor, the Papuan languages Makasae and Makalero have classifiers and couple these with other traces of Austronesian influence (Huber 2008, 2011). Papuan languages of the Bird's Head with classifiers include Mpur (Odé 2002), Abun (Berry and Berry 1999), Tehit (Flassy 1991), Maybrat (Dol 1999), Sougb (Reesink 2002b) and Hatam (Reesink

⁸ Iwam, Abau (in the East Sepik province), Chambri, Wogamusin, Chenapian (in the Lower Sepik), Angave, Tanae (in the Gulf Province), Folopa (in the Highlands), and Wantoat, Awará (in the Morobe province).

⁹ See Klamer, Reesink and Van Staden 2008; Klamer and Ewing 2010, Holton and Klamer, to appear, and references cited there.

1999), and these languages combine classifiers with Austronesian influence in e.g. word order, pronouns, numerals and lexicon (Voorhoeve 1989). An example from Mpur is (6):

Mpur (Odé 2002:83)

- (6) *Jan bik denur*
 house CLF three
 ‘Three houses.’

Before the advent of Indonesian, centuries of contact existed between Papuan speakers of the Bird’s Head and surrounding Austronesian speakers (the Wandamen in the south-east, people from the Raja Ampat islands in the south, and the Biak-Numfor in the north). Austronesian languages such as Biak were used as languages of wider communication in and around the Bird’s Head (van den Heuvel 2006), and it is very likely that this led to the diffusion of Austronesian features in the local Papuan languages (cf. Reesink 2002a:25-26). In Halmahera, the languages Tidore (Van Staden 2000:166-167) and Tobelo (Holton, this volume) also have classifiers. Old Austronesian loans are found throughout North Halmaheran Papuan languages suggest a very long period of contact between Papuan and Austronesian languages (Voorhoeve 1994).

In short, while numeral classifiers are extremely rare in Papuan languages generally, and do not occur in areal and/or genealogical clusters anywhere in mainland and eastern New Guinea, we do find them in Papuan languages spoken in those areas of Indonesia where Austronesian-Papuan contact is, or has been intense. I do not think this is accidental: it strongly suggests that Papuan classifiers developed under contact with Austronesian.

3.1 Numeral classifiers in Austronesian languages

In section 2 we saw that there is no evidence that proto-Alor-Pantar had one or more classifiers. We thus have to assume that classifiers in the individual Alor-Pantar languages developed after the proto-language split up.

In addition, the areal patterns discussed in the previous section suggest that this development was enhanced, or reinforced, by contact with Austronesian classifier languages. I propose that there have been at least two layers of contact: one very recent, and one presumably much more ancient.

The very recent contact involves Indonesian, the national language of Indonesia. As today’s language of interethnic communication, media and education, Indonesian is spoken as a second language by virtually everyone on Pantar and Alor. It also is the first language of an increasing number of children. The dominant role of Indonesian is a relatively recent phenomenon that started after the 1960’s, roughly correlating with the increasing number of Indonesian primary schools established in rural areas.

In this context it must be noted that Malay, which was the trade language in many parts of eastern Indonesia in pre-colonial times and the *lingua franca* of the Dutch East Indies, did not play a significant role in the history of the languages of Alor and Pantar. Alor and Pantar were under very remote Portuguese control until 1860, and Dutch colonial influence only became apparent in the first decades of the 20th century. In the Dutch era, few speakers of the local Papuan languages went to school and learned Malay (Klamer 2010:14 and references cited there). There is no evidence whatsoever that a Malay variety was used as a *lingua franca* across the islands in the past. Indeed, the local Malay variety that is currently spoken in the major town of Kalabahi is clearly based on the Malay variety that developed in the provincial capital of Kupang on Timor (Jacob and Grimes 2003) and has been introduced during the last few decades.

As today's most dominant language in the archipelago, Indonesian is influencing the local Papuan languages of Alor Pantar in many ways. Indonesian has a set of sortal classifiers that are obligatory in numeral contexts. Of these, the classifier *buah*, which is derived from a noun meaning 'fruit', is the "most general classifier [which] has almost lost any semantic, conceptual content" (Hopper 1986:323) and "classifies things that do not have definite types and shapes" (Chung 2010:553).¹⁰ An illustration is shown in (7).

Indonesian

- (7) *dua* ***buah*** *mobil*
 two CLF car
 'Two cars'

In this respect, Indonesian *buah* is similar to the general classifiers in Teiwa (*bag*), Western Pantar (*bina*), Adang (*pa'*) and Kamang (*uh*). It is very likely that recent, ongoing and intensive contact with Indonesian has been the driving force behind a development where these indigeneous lexemes became used as general classifiers over time. However, it must be emphasized that this is not a case of borrowing an Indonesian lexeme: neither form nor semantics of *buah* are copied in any of the languages. Indonesian *buah* derives from 'fruit', and as a general numeral classifier it classifies objects and fruits, not animals. In contrast, the Teiwa, Western Pantar, Adang and Kamang general classifiers do not derive from 'fruit' and do not generally classify fruits. The only feature shared by Indonesian *buah* and the general classifiers in these four languages is their *general* classifying function. This function may have been copied from Indonesian onto the various different lexemes in the local languages. In other words, the general Indonesian classifier *buah* may have driven or enhanced the development of general classifiers in the local Papuan languages of Alor and Pantar. Recent contact with Indonesian has thus lead to the development of general classifiers.

However, the numeral classifiers in Alor-Pantar languages were already in existence before the speakers got in contact with Indonesian. If the classifier sets developed due to contact with Austronesian classifier languages, as proposed above, this contact must have been from an earlier, more ancient date.

Classifiers are attested throughout the Austronesian-speaking world outside of Taiwan, that is, in the Malayo-Polynesian languages.¹¹ Examples include the Vietnam language Cham, Moken/Moklen on the Tai-Malay Peninsula, Belait in northern Borneo, Nias west of Sumatra, Mori Bawah in Sulawesi and Sama Bajau in the southern Philippines and eastern Indonesia (cf. Blust 2009:282-283, Himmelmann 2005: 173). Reflexes of the numeral classifier reconstructed for proto-Malayo-Polynesian **buaq* (Blust 2009:289) are found as general classifiers across the entire family, right down to the Oceanic subgroup.

We have good evidence that the Alor-Pantar languages have been in contact with Austronesian languages since the latter arrived in the area some 3000 years ago (Pawley 2005:100), as Austronesian loans have been reconstructed back to proto-Alor-Pantar (Holton et. al. 2012:114). Also, there are various types of Austronesian (but demonstrably non-Malay, non-Indonesian) lexical influences in the numeral systems of some of the Alor-Pantar languages (Schapper and Klamer 2014).

¹⁰ *Buah* is a reflex of the proto-Malayo-Polynesian noun **buaq* 'fruit' discussed below.

¹¹ All the Austronesian languages spoken outside of Taiwan, including those of Indonesia and the Oceanic languages, belong to the Malayo-Polynesian subgroup. When I use the term Austronesian here, it refers to "the Malayo-Polynesian subgroup of Austronesian".

Obvious candidates for contact with Alor-Pantar languages are (the predecessors of) the Austronesian languages spoken in the vicinity of Alor and Pantar, in eastern Indonesia. So it is useful to chart the occurrence of numeral classifiers in Austronesian languages spoken in eastern Indonesia. The number of languages spoken in eastern Indonesia may be estimated at 200-250 – a figure that is vague for lack of data (Hammarström and Nordhoff 2012), and we currently have grammatical information on some 25-30 of these languages (Klamer 2012). A cursory inspection of the existing grammars showed that most of them have numeral classifiers; examples are given in Table 7, and charted in Figure 4.

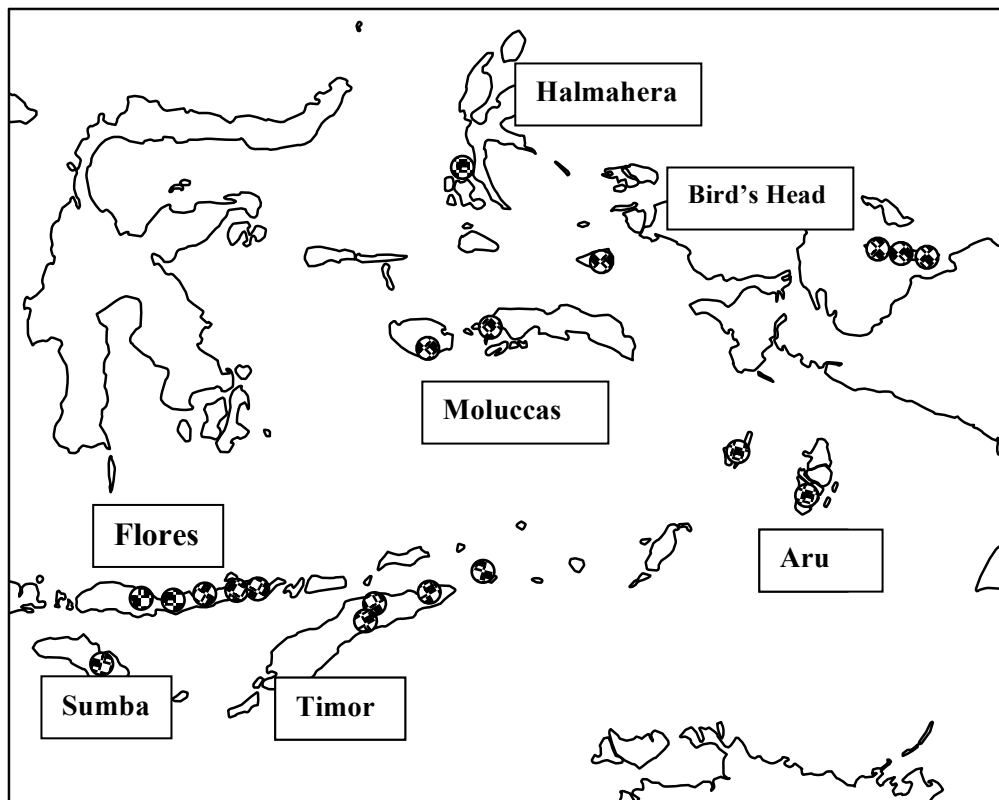


Figure 4. The Austronesian languages with numeral classifiers in eastern Indonesia from Table 7, represented as speckled dots

Table 7. Austronesian languages with numeral classifiers in eastern Indonesia

Region/island	Language	Source
Flores	Manggarai	Burger 1946
	Rongga	Arka 2008
	Kéo	Baird 2001, 2002
	Sika	Arndt 1931
	Hewa	Hanna Fricke (fieldnotes, 2014)
Timor	Tetun Fehan	Williams-van Klinken 1999
	Tetun Dili	Williams-Van Klinken et. al. 2002
	Waimaha	Hull 2002
	Leti	Van Engelenhoven 2004
Aru	Dobel	Hughes 2000
	Kei	Geurtjens 1921
Moluccas	Buru	Grimes 1991
	Larike	Laidig and Laidig 1995
Halmahera	Taba	Bowden 2001

Region/island	Language	Source
Bird's Head's region	Magéy Matbat	Remijsen 2010
	Kurudu	Anceaux 1961
	Ambai	Silzer 1983
	Munggui	Anceaux 1961
Sumba	Kambera	Klamer 1998

Illustrations of the classifiers found in the Austronesian languages of eastern Indonesia are given in Table 8-14. For reasons of space, I only present one language per island. Reflexes of the proto-Malayo-Polynesian form **buaq* 'fruit' are printed in bold.

Table 8. Numeral classifiers in Rongga (Flores) (Arka 2008)

Form	Meaning	Classifies
<i>mori</i>	'God'	humans
<i>ata</i>	'person'	humans
<i>esa</i>	(none)	general classifier for everything but humans animals and trees, including abstract objects (house, clan, idea, book title)
<i>eko</i>	'tail'	animals
<i>pu'u</i>	('pole?')	living standing trees
<i>li'e</i>	'fruit'	fruit-like three dimensional objects
<i>kura</i>	'bunch'	palm fruits such as coconuts
<i>wole</i>	'bunch'	banana, corn, rice
<i>kise</i>	'grain'	corn, tooth
<i>toko</i>	'log of wood'	non-flat, round, long objects (pipe, bamboo tube, sarong)
<i>mbi'i</i>	(none)	flat, thin, rigid, long objects (plank, board)
<i>nolo</i>	(none)	flat, thin, bendable, foldable objects (bamboo rope, shirt)

Table 9. Numeral classifiers in Tetun Fehan (Timor) (Van Klinken 1999:104-105)

Form	Meaning	Classifies
<i>na'in</i>	'noble', 'owner'	humans
<i>matan</i>	'source', 'eye'	large domestic animals that are eaten (pig, buffalo)
<i>lolon</i>	'trunk'	long objects (candle, fish, rib, woven cloth)
<i>tahan</i>	'leaf'	thin flat objects (clothing, betel pepper, book)
<i>fuan</i>	'fruit', 'heart'	whole roundish objects (betel nut, coconut, cabbage, egg, sea shell)
<i>musan</i>	'seed'	very small round objects (tablet)

Table 10. Numeral classifiers in Dobel (Aru) (Hughes 2000:158-159)

Form	Meaning	Classifies
<i>fatin</i>	'body'	humans, trees
<i>fusi</i>	'fruit'	fruits, other (e.g. human shouts, dog barks)
<i>yafir</i>	'shaft'	long, pole-shaped objects (trousers)
<i>rakwin</i>	'leaf'	thin flat objects (leaf, cloth)
<i>fa'il</i>	(none)	thicker flat objects (plank, sago filter)
<i>kwasir</i>	(none)	boats and villages

Table 11. Numeral classifiers in Buru (Moluccas) (Grimes 1991:306-310)¹²

Form	Meaning	Classifies
<i>geba</i>	‘person’	humans
<i>ihēr</i>	‘thing’	general classifier used if other classifiers are unknown or unimportant
<i>fatan</i>	‘trunk’	long, large and round (tree trunks, waves, bodies of (dead) animals)
<i>kisen</i>	‘bald’	live pigs and chicken
<i>isin</i>	‘content’	tubers
<i>fuan</i>	‘fruit’	any bulbous shaped thing (fruit, vegetables, wok, bread, pearl)
<i>somon</i>	‘part’	whole clothes
<i>lahin</i>	‘root’	trees
<i>fatun</i>	‘stone’	rock-like objects
<i>tian</i>	‘belly’	corn
<i>kasen</i>	‘section’	things that have natural divisions (bamboo, language)
<i>somon</i>	‘part’	(no information)
<i>engan</i>	‘piece’	meat, cloth (implies usability of piece)
<i>lafan</i>	‘sheet’	thin, flat things (leaves, paper, cloth)
<i>wangan</i>	‘digit’	of definable length (finger, short strip of bamboo)
<i>walan</i>	‘strip’	objects with salient feature of length (hair, strip of pandanus leaf for weaving)
<i>walen</i>	‘pole’	bamboo
<i>turen</i>	‘short’	short length of wood, bamboo
<i>kodon</i>	‘straw’ (?)	house structures
<i>kihan</i>	‘seed’	small seeds, rice, sand
<i>folin</i>	‘stalk’	bananas
<i>dahan</i>	‘hand’	bananas
<i>pongon</i>	‘cluster’	betelnut, grapes

Table 12. Numeral classifiers in Taba (Halmahera) (Bowden 2001:252-263)¹³

Form	Meaning	Classifies
<i>p-</i> ¹⁴	(none)	general classifier attached to numerals 1-9
<i>i-, sis=</i>	(none)	animal classifier attached to numerals 1-9, day/sun, month/moon
<i>mat=</i>	(none)	humans
<i>mot=</i>	(none)	small, square flat thin and cut objects
<i>wato-</i>	(none)	small oblong shaped objects (not cut) (pen, match)
<i>hola</i>	(none)	pieces of wood, sticks
<i>luklik</i>	‘to roll s.th.’	rolled up objects (cigarette)
<i>ai</i>	‘tree’	trees
<i>awa</i>	‘stalk’	bunches of fruits growing together on a stalk
<i>ising</i>	‘hand’	hands of bananas
<i>kop</i>	‘grain’	grains of rice and corn

¹² I have included here the Buru forms that are classified as sortal and/or mensural, and excluded those that are classified separately as mensural classifiers by Grimes.

¹³ The boundary between sortal and mensural classifiers in Taba is unclear. I have excluded here the forms that seemed transparently mensural.

¹⁴ Prefix to numerals, probably derived from proto-form **buaq*, see the discussion below.

boka ‘skewer’ things skewered on a stick

Table 13. Numeral classifiers in Magey Matbat (Misool, W of Bird’s Head) (Remijsen 2010:287-290)

Form	Meaning	Classifies
<i>nun-</i>	(none)	animates
<i>ha-</i>	(none)	boats and houses
<i>di-</i>	(none)	long sharp objects
<i>pa-</i>	(none)	round objects, often fruits
<i>ta-</i>	(none)	sago biscuits
<i>i-</i>	(none)	general classifier for everything else including abstract nouns

Table 14. Numeral classifiers in Kampera (Sumba) (Klamer 1998:139, Onvlee 1984)

Form	Meaning	Classifies
<i>tau</i>	‘person’	humans
<i>iu/ngiu</i>	(none)	animals
<i>wua/mbua</i>	fruit	general classifier for three dimensional and abstract objects (fruit, cup, house, month)
<i>pungu/mbungu</i>	pole	oblong objects (stick, needle)
<i>wàla/mbàla</i>	leaf	flat thin objects (cloth, paper)

The data presented in Table 8-14 contrast with the view expressed by Blust (2009: 282), who remarks that “[classifiers] are reported sporadically in eastern Indonesia”. Indeed, numeral classifiers are almost univesal in the Austronesian languages of eastern Indonesia. And those languages that do not have numeral classifiers as independent words often have prefixes to numerals that can be shown to historically derive from one or more classifier(s). Illustrations include the prefix *p-* in Taba (Table 12), probably derived from proto-Austronesian/Malayo-Polynesian **buaq* (Bowden 2001: 244), and the prefix *pa-* in Magey Matbat (Table 13), illustrated in (8), which is likely to be of the same source:

Magey Matbat (Remijsen 2010: 287)

- (8) *Ho^lng pa-lu^l*
mango CLF-two
‘two mangos’

A further illustration of a grammaticalised numeral classifier is the (synchronically meaningless) prefix *v(ò)-* in Leti, which is found only attached to the numerals ‘two’ to ‘nine’, e.g. *vò-rua* ‘two’, *vò-telu* ‘three’ (Van Engelenhoven 2004:164). This prefix is also a reflex of proto-Malayo-Polynesian **buaq*, compare Leti *vua/vua* ‘fruit’ (Van Engelenhoven 2004:447).¹⁵

¹⁵ Similar constructions where classifiers have become meaningless affixes attached to numerals are reported for Oceanic, where “generally, the numeral and classifier are bound to each other in one or the other order” (Lynch, Ross and Crowley 2002:73).

In short, numeral classifiers are very common in the Austronesian languages of eastern Indonesia, and while they show a lot of variation in shape and categorical functions, reflexes of pMP **buaq* are found across the region.

Classifiers are also attested widely in the Oceanic subgroup of Austronesian. A number of classifiers have been reconstructed for proto-Oceanic (pOC), including the general classifier **puaq* literally ‘fruit’ in (9). Other reconstructed forms are the classifier for animate beings **mwane*, a classifier for wooden or elongated objects (**kaiu*) and one for persons (**tau*) (Lynch et al. 2007:74). The Oceanic languages with classifiers include those listed in (10) below (taken from Lynch et al. 2002:73-74). Major subgrouping information is included in brackets.

Proto-Oceanic (Lynch et.al. 2002:73)

- (9) *ta* ***puaq*** *tolu* *a* *niuR*
 ART CLF three ART coconut
 ‘Three coconuts’

- (10) The Admiralties family (a primary subgroup of Oceanic languages)
 The Kilivila family (< Western Oceanic linkage)
 Sudest (Papuan Tip < Western Oceanic linkage)
 The North Bougainville linkage (< Meso-Melanesian linkage < W Oceanic linkage)
 The Cristobal-Malaitan languages (< SE Solomonian family < CE Oceanic linkage)
 The Nuclear Micronesian family (< CE Oceanic linkage)
 The languages in New Caledonia (<S Oceanic linkage < CE Oceanic linkage)
 The Polynesian languages (<Central Pacific linkage < CE Oceanic linkage)

All this goes to show that classifier languages are found across the three primary subgroups of Oceanic, as well as across the subgroups *within* each of these subgroups, and that they are reconstructed for proto-Oceanic. Classifiers must have been quite commonly used in proto-Oceanic. For instance, Lynch et. al (2002:74) mention that a word like **niuR* ‘coconut’ depended on a classifier for its disambiguation, as the meaning of **niuR* itself embraced the notion of the tree, its fruit, and the contents of the fruit.

Since Proto Malayo-Polynesian had one classifier (**buaq*) and Proto Oceanic had four (**puaq*, **mwane*, **kaiu*, **tau*; see above), it is safe to assume that the Malayo-Polynesian (Central Eastern Malayo Polynesian, Blust 2009) languages of eastern Indonesia today had at least one classifier, and possibly more, since they constitute the link between Malayo-Polynesian and Oceanic, shown in Figure 5.

In sum, numeral classifiers are found all across the Austronesian languages spoken outside of Taiwan; they are heavily attested throughout eastern Indonesia; and they have been reconstructed for the Oceanic subgroup. On the other hand, they are not typically found in Papuan languages, except for those that are spoken in the vicinity of Austronesian languages, including the languages of Alor and Pantar. This makes it likely that contact with Austronesian languages has contributed to the development of numeral classifiers in the Papuan languages of Alor and Pantar.

Note again that the contact did not involve a diffusion of lexemes: no similarity in shape or semantics exists between classifiers in Alor-Pantar languages and classifiers of Austronesian languages in the area. In particular, reflexes of the reconstructed proto Malayo-Polynesian form **buaq*, which are found throughout the Austronesian family, do not occur in any of the Alor-Pantar languages.

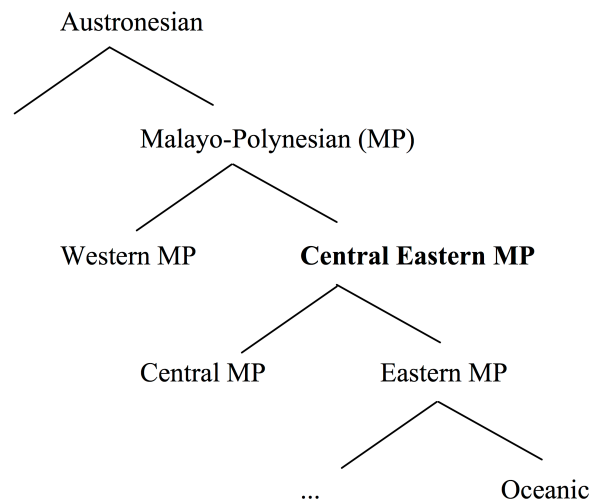


Figure 5. Austronesian language family and its main branches

4 Conclusions

Survey data on Alor-Pantar numeral classifiers indicate that numeral classifiers are very common in Alor and Pantar, but have variable shapes and origins, and make very different semantic classifications. No classifier is reconstructable for proto-Alor-Pantar. This in turn suggests that the classifier systems found in the family developed after the split up of proto-Alor-Pantar. This is not a surprising finding, as it is well known that classifier sets are highly volatile and always develop out of other lexical classes such as nouns.

However, from a Papuan point of view, the development of sets of numeral classifiers seems rather unusual, as numeral classifiers are extremely rare in Papuan languages generally. They do not occur in any areal and/or genealogical cluster of Papuan languages, *except* for three areas in eastern Indonesia: the Bird's Head, Halmahera and Timor-Alor-Pantar. These are exactly the three Indonesian regions known to have undergone long term Austronesian-Papuan contact, resulting in diffusion of structural features.

Classifiers are typical for Austronesian languages, and the Austronesian languages in eastern Indonesia almost universally have them, so it seems plausible that the development of classifiers in the Alor-Pantar languages was triggered or enhanced by (ancient) Austronesian influence.

In addition, recent and intensive contact with Indonesian may have lead to the development of general classifiers in a good number of Alor-Pantar languages, as functional copies of the Indonesian general classifier *buah*.

The contact did not involve any borrowing of lexemes: no similarity in shape or semantics exists between classifiers in Alor-Pantar languages and classifiers of Austronesian languages in the area. In particular, no reflexes of the reconstructed proto-Malayo-Polynesian form **buaq*, which are found throughout the Austronesian family, are attested in the Alor-Pantar languages surveyed here. Neither has the grammatical structure of Austronesian numeral NPs been copied. In Austronesian NPs, the classifier follows the numeral, while the position of the noun varies, thus we find [NUMERAL - CLASSIFIER - NOUN] (as in Indonesian *dua buah rumah* 'two CLF houses') but also [NOUN - NUMERAL - CLASSIFIER] (as in colloquial Malay *rumah dua buah* 'houses two CLF') (Blust 2009: 283-

284). In contrast, in numeral NPs in the Alor-Pantar languages the classifier precedes the numeral: [NOUN - CLASSIFIER - NUMERAL].¹⁶

This suggests that the Alor-Pantar classifiers indeed constitute an independent development. What speakers may adopted from Austronesian, however, is the propensity to reanalyse lexemes they already had at their disposal (such as part-of-whole nouns) and grammaticalise them as classifiers in numeral expressions.

By comparing the numeral classifier sets used in the Alor-Pantar languages with each other, as well as with classifier patterns of Papuan and Austronesian languages more generally, we have seen that most of the Alor-Pantar numeral classifiers developed out of nouns. The process of grammaticalization was not only internally motivated, but also modelled after functions found in Austronesian languages with numeral classifiers.

References

- Aikhenvald, Alexandra Y., 2000, *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press.
- 2006, *Classifiers and Noun Classes: Semantics*. In *Encyclopedia of Languages and Linguistics*, 463-471. Oxford UK: Elsevier.
- Aikhenvald, Alexandra Y., and Tonya Stebbins, 2007, *Languages of New Guinea*. In O. Miyaoka, O. Sakiyama and M.E. Krauss, eds. *The Vanishing Voices of the Pacific Rim*, 239-66. Oxford: Oxford University Press.
- Anceaux, J.C., 1961, *The linguistic situation in the islands of Yapen, Kurudu, Nau and Miosnum, New Guinea*. *Verhandelingen van het Koninklijk Instituut voor Taal-, Land-, en Volkenkunde* 35. The Hague: Nijhoff.
- Arka, I Wayan, 2008, *Classifiers in Rongga*. *Studies in Philippine Languages and Cultures* 18:1-17.
- Arndt, Paul, 1931, *Grammatik der Sika-Sprache*. Ende: Arnoldus.
- Baird, Louise, 2001, *Numerals and classifiers in Keo*. *Linguistika* 14:1-22.
- 2002, *A grammar of Kéo: An Austronesian language of East Nusantara*. Ph.D. thesis Australian National University.
- 2008, *A Grammar of Klon: A Non-Austronesian Language of Alor, Indonesia*. Canberra: Pacific Linguistics.
- Berry, Keith, and Christine Berry, 1999, *A Description of Abun*. Canberra: Pacific Linguistics.
- Blust, Robert, 2009, *The Austronesian languages*. Canberra: Pacific Linguistics.
- Bowden, John, 2001, *Taba: Description of a South Halmahera Austronesian language*. Canberra: Pacific Linguistics.
- Burger, P.A., 1946, *Voorlopige Manggaraise Spraakkunst*. *Bijdragen tot de Taal-, Land- en Volkenkunde* 103(1-2):15-265.
- Chung, Siaw-Fong, 2010, *Nominal classifier buah in Malay: A corpus-based study*. *Language and Linguistics* 11(3):553-577.

¹⁶ Note that example (8) from Austronesian Magey Matbat does not represent the proto-Austronesian structure. It is more likely a structure that is based on a Papuan substrate of the Raja Ampat islands.

- Dol, Philomena, 2007, *A Grammar of Maybrat: A language of the Bird's Head, Papua Province, Indonesia*. Canberra: Pacific Linguistics.
- van Engelenhoven, Aone, 2004, *Leti, a language of Southwest Maluku*. Leiden: KITLV Press.
- Flassy, Don Augusthinus Lamaech, 1991, *Grammar sketch of Tehit: A Toror language, the West Doberai Peninsula, New Guinea (Irian Jaya)*. M.A. Thesis Leiden University.
- Foley, William A., 1986, *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- 2000, The languages of New Guinea. *Annual Review of Anthropology* 29:357-404.
- Geurtjens, Hendrik, 1921, *Spraakleer der Keieesche taal*. *Verhandelingen van het Bataviaansch Genootschap voor Kunsten en Wetenschappen* 63.2. Albrecht and Martinus Nijhoff.
- Gil, David, 2011, Numeral classifiers. In Matthew Dryer and Martin Haspelmath, eds. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Available online at <http://wals.info/>. Accessed on 2013-02-11.
- Grimes, Charles E., 1991, *The Buru language of Eastern Indonesia*. Ph.D. thesis Australian National University.
- Haan, Johnson Welem, 2001, *A grammar of Adang: a Papuan language spoken on the Island of Alor, East Nusa Tenggara, Indonesia*. Ph.D. thesis University of Sydney.
- Hammarström, Harald, and Sebastian Nordhoff, 2012, The languages of Melanesia: Quantifying the level of coverage. In Nicholas Evans and Marian Klamer (eds.), *Melanesian languages on the edge of Asia*, *Language Documentation & Conservation Special Publication No. 5*, 13–33. < <http://nflrc.hawaii.edu/ldc/sp05/>>, <<http://hdl.handle.net/10125/4559>>.
- van den Heuvel, Wilco, 2006, *Biak: A description of an Austronesian language of Papua*. Ph.D. thesis Vrije Universiteit. Utrecht: LOT Publications.
- Himmelmann, Nikolaus, 2005, The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar and Nikolaus P. Himmelmann, eds. *The Austronesian languages of Asia and Madagascar*, 110-181. London and New York: Routledge.
- Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, Dmitry and Egorov, 2011, Automated dating of the world's language families based on lexical similarity. *Current Anthropology* 52(6):841-75.
- Holton, Gary, This volume, *Numeral classifiers and number in two Papuan outliers of East Nusantara*.
- To appear, *Sketch of Western Pantar (Lamma)*. In Antoinette Schapper, ed. *Papuan Languages of Timor-Alor-Pantar: Sketch grammars*. Berlin: De Gruyter Mouton.
- Holton, Gary, Marian Klamer, František Kratochvíl, Laura C. Robinson, and Antoinette Schapper, 2012, The historical relation of the Papuan languages of Alor and Pantar. *Oceanic Linguistics* 51(1):87-122.
- Holton, Gary and Marian Klamer. To appear. *The Papuan languages of East Nusantara*. In Bill Palmer, ed. *Oceania*. Berlin: Mouton De Gruyter.

- Hopper, Paul J., 1986, Some discourse functions of classifiers in Malay. In Colette G. Craig (ed.), *Noun Classes and Categorization*, 309-325. Amsterdam: John Benjamins.
- Huber, Juliette, 2008, *First steps towards a grammar of Makasae*. Munich: LINCOM Europa.
- 2011, *A grammar of Makalero: A Papuan language of East Timor*. Ph.D. thesis Leiden University. Utrecht: LOT Publications.
- Hughes, Jock, 2000, The Morphology of Dobel, Aru, with Special Reference to Reduplication. In Chuck E. Grimes, ed. *Spices from the East: Papers in Languages of Eastern Indonesia*, 131-180. Canberra: Pacific Linguistics.
- Hull, Geoffrey, 2002, *Waimaha*. Instituto Nacional de Linguística, Universidade Nacional de Timor Lorosa'e.
- Jacob, June and Charles E. Grimes, 2003, *Kamus Pengantar Bahasa Kupang*. Kupang: Artha Wacana Press.
- Klamer, Marian, 1998, *A grammar of Kambara*. Berlin: Mouton de Gruyter.
- 2010, *A grammar of Teiwa*. Berlin: Mouton de Gruyter.
- 2012, Tours of the past through the present of eastern Indonesia. In Frank Seifart, Geoffrey Haig, Nikolaus P. Himmelmann, Dagmar Jung, and Paul Trilsbeek, eds. *Potentials of Language Documentation: Methods, Analyses, and Utilization*, Language Documentation and Conservation Special Publication No. 6, 55-64. <<http://www.nflrc.hawaii.edu/ldc/sp03/>>
- 2014, The history of numeral classifiers in Teiwa (Papuan). In Gerrit J. Dimmendaal and Anne Storch, eds. *Number: Constructions and Semantics. Case studies from Africa, India, Amazonia & Oceania*, 135-166. Amsterdam: Benjamins.
- Klamer, Marian, Ger P. Reesink, and Miriam van Staden, 2008, East Nusantara as a linguistic area. In Pieter Muysken, ed. *From Linguistic Areas to Areal Linguistics*, 95-150. Amsterdam: John Benjamins.
- Klamer, Marian, and Michael Ewing, 2010, The languages of East Nusantara: an introduction. In Michael Ewing and Marian Klamer, eds. *East Nusantara: Typological and Areal Analyses*, 1-24. Canberra: Pacific Linguistics.
- Laidig, Wyn D., and Carol J. Laidig, 1995, A synopsis of Larike phonology and syntax. In Wyn D. Laidig, ed. *Descriptive Studies in Languages of Maluku*, 19-42. Jakarta: NUSA 2.
- Lynch, John, Malcolm Ross and Terry Crowley, eds. 2002, *The Oceanic Languages*. Richmond Surrey: Curzon Press.
- Lyons, John, 1977, *Linguistic Semantics: An introduction*. Cambridge: Cambridge University Press.
- Odé, Cecilia, 2002, A Sketch of Mpur. In Ger P. Reesink, ed. *Languages of the East Bird's Head*, 45-107. Canberra: Pacific Linguistics.
- Onvlee. Louis, 1984, *Kambaraas (Oost-Sumbaas)-Nederlands Woordenboek*. Dordrecht: Foris Publications.
- Pawley, Andrew K., 2005, The chequered career of the Trans New Guinea hypothesis. In Andrew Pawley, Robert Attenborough, Jack Golson and Robin Hide, eds. *Papuan Pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*, 67-107. Canberra: Pacific Linguistics.

- Reesink, Ger P., 1999, A grammar of Hatam, Irian Jaya, Indonesia. Canberra: Pacific Linguistics.
- 2002a, Languages of the Eastern Bird's Head (Pacific Linguistics 524). Canberra: Pacific Linguistics.
- 2002b, A grammar sketch of Sougb. In Ger P. Reesink, ed. Languages of the East Bird's Head, 181-275. Canberra: Pacific Linguistics.
- Remijnsen, Bert, 2010, Nouns and verbs in Magey Matbat. In Michael Ewing and Marian Klamer, eds. East Nusantara: Typological and Areal Analyses, 252-280. Canberra: Pacific Linguistics.
- Robinson, Laura and John W. Haan, To appear, Adang. In Antoinette Schapper, ed. Papuan Languages of Timor-Alor-Pantar: Sketch grammars. Berlin: De Gruyter Mouton.
- Robinson, Laura and Gary Holton, 2012, Internal classification of the Alor-Pantar language family using computational methods applied to the lexicon. *Language Dynamics and Change* 2(2):1-27.
- Schapper, Antoinette, 2010, Abui Numerals Questionnaire. Unpublished ms. Leiden University.
- To appear. Kamang. In Antoinette Schapper, ed. Papuan Languages of Timor-Alor-Pantar: Sketch grammars. Berlin: De Gruyter Mouton.
- Schapper, Antoinette, Juliette Huber and Aone van Engelenhoven, 2012, *The historical relations of the Papuan languages of Timor and Kisar*. *Language and Linguistics in Melanesia*, 192-240.
- Schapper, Antoinette and Marian Klamer, 2014, Numeral systems in the Alor-Pantar languages. In Marian Klamer, ed. *The Alor-Pantar languages: History and Typology*. Berlin: Language Science Press, < <http://langsci-press.org/>>.
- Shimojo, Mitsuaki, 1997, The Role of the General Category in the Maintenance of Numeral Classifier Systems: The case of *tsu* and *ko* in Japanese. *Linguistics* 35, 4, 705-733.
- Silzer, Peter James, 1983, Ambai, an Austronesian language of Irian Jaya, Indonesia. Ph.D. thesis Australian National University.
- van Staden, Miriam, 2000, Tidore: A Linguistic Description of a Language of the North Moluccas. Ph.D. thesis Leiden University.
- Voorhoeve, C. L., 1989, The masked bird: Linguistic relations in the Bird's Head area. In P. Haenen and J. Pouwer, eds. *Peoples on the Move: Current Themes of Anthropological Research in New Guinea*, 78-101. Nijmegen: University of Nijmegen.
- 1994, Contact-induced language change in the non-Austronesian languages in the north Moluccas, Indonesia. In T. Dutton and D.T. Tryon, eds. *Language Contact and Change in the Austronesian World*, 649-74. Berlin: Mouton de Gruyter.
- Williams-Van Klinken, Catharina, 1999, A grammar of the Fehan dialect of Tetun, an Austronesian language of West Timor. Canberra: Pacific Linguistics.
- Williams-Van Klinken, Catharina, John Hajek and Rachel Nordlinger, 2002, Tetun Dili: a grammar of an East Timorese language. Canberra: Pacific Linguistics.
- Zubin, David A. and Mitsuaki Shimojo, 1993, How 'general' are General Classifiers? With special reference to *ko* and *tsu* in Japanese. *Proceedings of the Nineteenth Annual*

Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals, 490-502.