

[Appeared as: Klamer, Marian. 2014. The history of Numeral classifiers in Teiw (Papuan). In Gerrit J. Dimmendaal and Anne Storch (eds.), *Number: Constructions and Semantics. Case studies from Africa, India, Amazonia & Oceania*. Amsterdam: Benjamins, 135-166.]

## **The history of Numeral classifiers in Teiwa (Papuan)**

Marian Klamer

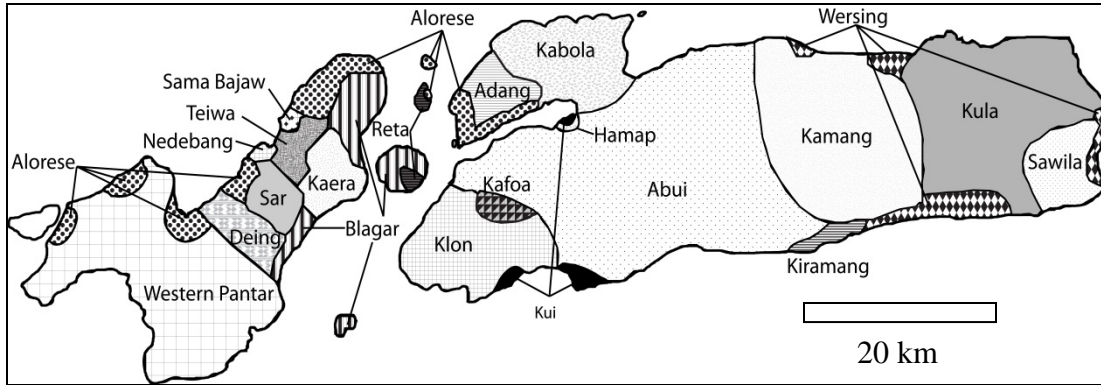
### **Abstract**

*The Papuan language Teiwa has a small set of sortal numeral classifiers: one human classifier, three fruit classifiers, and a general classifier. The classifiers vary widely in function and distribution, and it is argued that it is unlikely that they have inherited. Instead, it is proposed that Teiwa inherited a class of part-of-whole nouns from which certain members were recruited to be grammaticalised into numeral classifiers through the reanalysis of ambiguous structures. Two more factors have enhanced this development. First, the existence of number neutral nouns in Teiwa implies that nouns must be individuated before they can be counted, a function fulfilled by numeral classifiers. Second, areal pressure from Austronesian classifier languages has reinforced the development of classifiers. Connecting the numeral classifier system of Teiwa with those of its sister languages and the wider linguistic context of eastern Indonesia, we can thus identify structural, semantic and areal factors playing a role when classifiers are born into a language.*

**Keywords:** Papuan, sortal numeral classifiers, part-of-whole nouns, grammaticalisation, Papuan- Austronesian contact

## 0. Introduction<sup>1</sup>

Teiwa is a Papuan (‘non-Austronesian’)<sup>2</sup> language with 4,000 speakers who live in the villages of Madar and Lebang, in the North-West of Pantar island, in eastern Indonesia, see Maps 1 and 2. Klamer 2010 is a grammar of the language.<sup>3</sup>

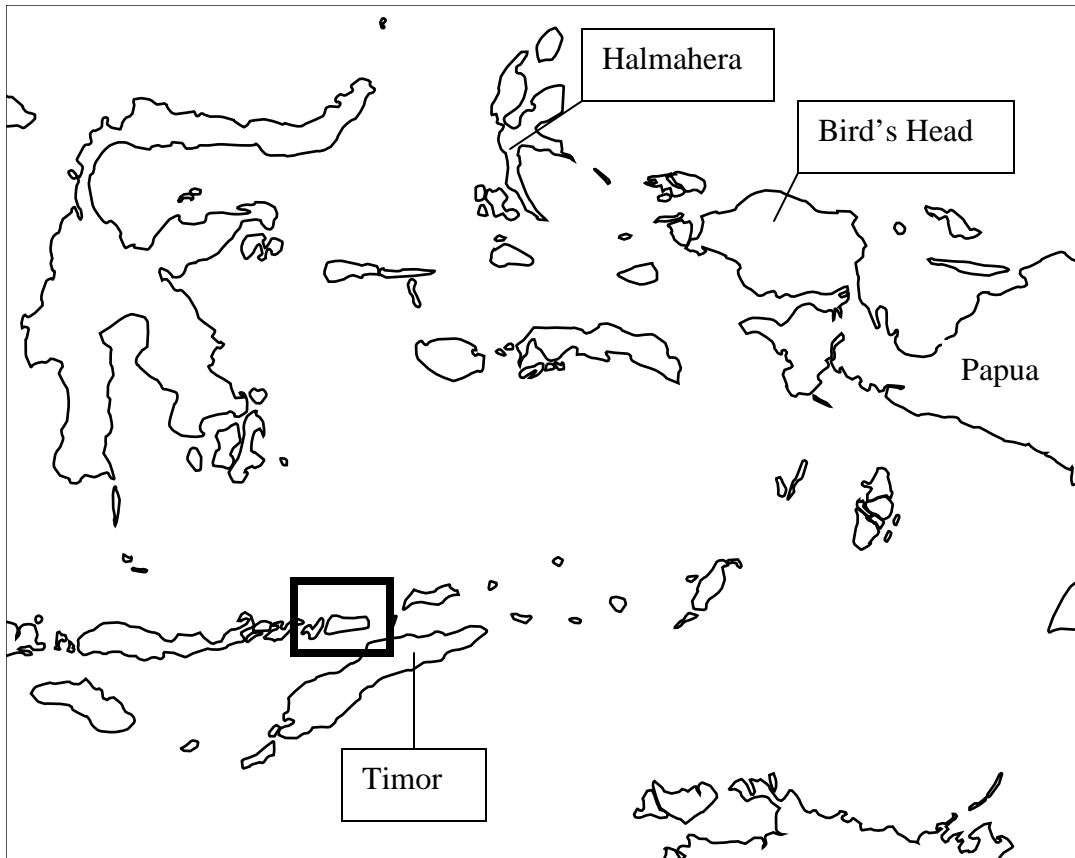


Map 1. The languages on the islands of Pantar (left) and Alor (right).

<sup>1</sup> I would like to thank Sascha Aikhenvald and an anonymous reviewer for comments on an earlier draft. The abbreviation IND indicates an Indonesian loan word. In Teiwa orthography a hyphen represents a glottal stop, *q* a uvular stop and *x* a voiceless pharyngeal fricative.

<sup>2</sup> The term ‘Papuan’ refers to unrelated families of languages spoken in New Guinea or its vicinities.

<sup>3</sup> The data presented in this paper were collected after the grammar was published. Where there are discrepancies between Klamer 2010 and the current paper, the latter prevails.



Map 2. Location of Alor and Pantar in Eastern Indonesia.

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 2005). In this paper, the term ‘classifier’ refers to sortal numeral classifiers; mensural classifiers will not feature in the present discussion.

Teiwa is a member of the Timor Alor Pantar (TAP) family. This family comprises ~25 Papuan languages that are spoken on Timor, Alor, Pantar, and islets in their vicinity. The TAP family branches into the Alor Pantar (AP) group, with ~20 members (Holton et al. 2012) and the Timor group with 5 members (Schapper et al. 2012). Teiwa is a member of

the Alor Pantar (AP) subgroup. There is evidence that the Alor Pantar languages originate from the Straits region between Pantar and Alor (Robinson & Holton 2012).

The ‘Papuan’ character of the TAP languages has long been recognized in the literature. Beginning with Wurm, Voorhoeve, and McElhanon (1975), most authors have assumed, mainly on structural evidence, that the TAP family belongs to the putative Trans-New Guinea family. In the absence of supporting lexical evidence, Holton et al. (2012) instead propose that the TAP group should be considered a distinct family, unrelated to Trans-New Guinea, and this is 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.

In this paper I first present a description of the Teiwa classifiers (section 1). Then I address the question where they could have originated from, and propose a grammaticalisation path (section 2). Next I discuss possible motivations for the development. One is that Teiwa has number neutral nouns and use classifiers to individuate nouns in counting constructions (section 3). Another force in the genesis of classifiers is intensive contact with classifier languages (section 4). A summary is presented in section 5.

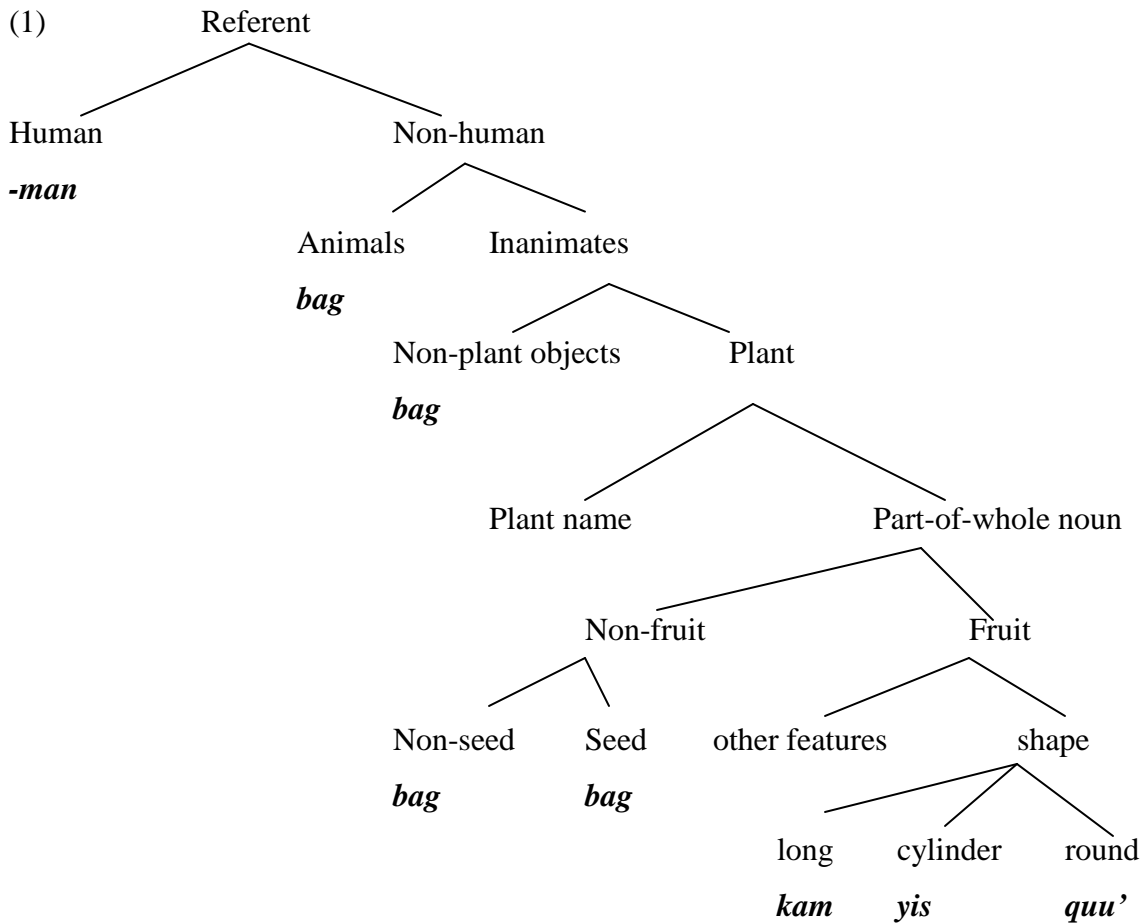
## 1. Overview of Teiwa classifiers

Teiwa has five numeral classifiers which come in three types, see Table 1 and the diagram in (1).<sup>4</sup>

Table 1. Numeral classifiers in Teiwa

Type	Form	Gloss	Classifies
Human	<i>-man</i>	CLF.HUM	humans
Fruit shape	<i>kam</i>	FRUIT:LONG	long fruits, e.g. tamarind, banana
	<i>yis</i>	FRUIT:CYLINDRICAL	cylindrical fruits and tubers, e.g. taro, cassava
	<i>quu'</i>	FRUIT:ROUND	round fruits: mango, papaya, lemon, pumpkin, coconut
General	<i>bag</i>	CLF, originally ‘seed’	everything except humans and fruits

<sup>4</sup> Abbreviations: ART = article, CLF = classifier, EXCL = exclusive, GEN = genitive, HUM = human, INCL = inclusive, IND = Indonesian loan, PL = plural, SG = singular.



The only nouns that always take a classifier are fruits. Humans take an optional classifier, and animals and inanimate non-plant objects can optionally occur with the general classifier *bag*. Teiwa has no dedicated classifier for animals nor for inanimate objects. Classifiers always occur in between the noun and the numeral. Below I first discuss the properties of the human classifier (section 1.1), then the fruit shape classifiers (section 1.2) followed by the general classifier *bag* that developed from a noun ‘seed’ (section 1.3). Section 1.4 presents a summary.

### 1.1. The human classifier –man

When humans are counted in Teiwa, a pronoun is used that is constructed out of a classifier base *–man* (or its metathesized form *–nam*) and a person-marking prefix. The paradigm of human classifier pronouns is presented in Table 2.

The person-marking prefixes that attach to the human classifier are identical to the short form of plural subject (S/A) pronouns, see Table 3. Teiwa short and long subject pronouns have the same grammatical function and occupy the same position in the clause, but the long forms typically express contrastive focus while the short forms are never used in such contexts (Klamer 2010:165-166).

Table 2. Human classifier pronouns in Teiwa

1PL.EXCL	<i>ni-man</i>
1PL.INCL	<i>pi-man</i>
2PL	<i>yi-man</i>
3PL	<i>i-man</i>

Table 3. Long and short subject pronouns in Teiwa

1SG	<i>na, na'an</i>	1PL.EXCL	<i>ni, ni'in</i>
		1PL.INCL	<i>pi, pi'in</i>
2SG	<i>ha, ha'an</i>	2PL	<i>yi, yi'in</i>
3SG	<i>a, a'an</i>	3PL	<i>i, iman</i> <sup>5</sup>
		3PL.ELSEWHERE	<i>i, i'in</i>

In NPs that enumerate humans and are uttered in isolation, for example, as response to a question ‘How many women were there?’, the classifier pronoun is obligatory, as shown in (2a-b). If the NP occurs in a different kind of context the human classifier can be omitted, as indicated by the brackets in (2c).

<sup>5</sup> The long pronoun *iman* ‘3PL’ is identical to the third person classifier pronoun. *Iman* has replaced the morphologically regular form *i'in*, which developed a more specific function ‘3PL.ELSEWHERE’ to refer to people that are at a place or time different from the speaker (Klamer 2010:79-80).

- (2) a. *Eqar i-nam yusan*  
 woman 3PL-CLF.HUM five  
 ‘(There were) five women’
- b. \* *Uy yusan*  
 person five
- c. *Uy (i-nam) yusan maraqai Qalambas ma gi.*  
 person 3PL-CLF.HUM five up Kalabahi come go  
 ‘Five people go up to Kalabahi.’

The human classifier pronoun is obligatory in questions with the interrogative quantifier *yiran* ‘how many/much’, compare (3a-b).

- (3) a. *Uy i-man yiran?*  
 person 3PL-CLF.HUM how.many/much  
 ‘How many people?’
- b. \* *Uy yiran?*  
 person how.many/much

While the Teiwa free pronouns can encode human referents, they cannot be used when humans are counted. For example, the long pronoun *pi'in* ‘1PL.INCL’ cannot be used in (4a); instead the human classifier pronoun *pi-man* is used, as in (4b).

- (4) a. \* *Pi'in ut ina.*  
 1PL.INCL four eat
- b. *Pi-man ut ina.*  
 1PL.INCL-CLF.HUM four eat  
 ‘We four eat’

The human classifier can be used for lower numerals as well as high ones, compare (5) with (2c). The classifier may optionally be omitted as indicated by the brackets.

- (5) *Uy* (*i-nam*) *ratu* *nuk* *aria'-an* *gereja*.  
 person 3PL-CLF.HUM hundred one arrive-REAL church  
 'One hundred people came to church.'

A human classifier cannot co-occur with a quantifier, as illustrated in (6).

- (6) a. *Eqar* *dum*  
 woman many/much  
 'many women'
- b. \* *Eqar* *i-man* *dum*  
 woman 3PL-CLF.HUM many/much

In sum, a human classifier pronoun which combines a plural pronominal prefix with *-man* is used to enumerate human referents. The classifier pronoun is obligatory in questions with the interrogative quantifier *yiran* 'how many/much', as well as in answers to such questions. When an enumerated noun occurs in a larger clausal context, for example with an activity predicate, the human classifier may be omitted.



## 1.2. The fruit shape classifiers

When and tubers are counted, classifiers must be used to classify them according to their shape. The fruit classifiers classify fruits and tubers according to their shape, as in Table 4.

Table 4. Fruit/tuber classifiers in Teiwa (repeated from Table 1)

Form	Gloss	Classifies
<i>kam</i>	FRUIT:LONG	long fruits, e.g. tamarind, banana
<i>yis</i>	FRUIT:CYLINDRICAL	cylindrical fruits and tubers, e.g. taro, cassava
<i>quu'</i>	FRUIT:ROUND	round fruits: mango, papaya, lemon, pumpkin, coconut

The fruit shape classifiers combine with generic plant nouns. In this way they create an expression to refer to fruits of that plant. For example, in (7a), the classifier *quu'* ‘CLF.FRUIT:ROUND’ combines with *wou*, a generic word for everything related to ‘mango’ – the tree, branches, leaves, fruits, or flowers. To capture this generic function, the noun *wou* is glossed as ‘mango-hood’, because, unlike what an English gloss ‘mango’ would suggest, *wou* by itself cannot refer to the fruit: Teiwa *wou* is never referential on its own, as shown by the meaningless utterance in (7b). To refer to the fruit, *wou* must combine with the fruit classifier *quu'*, and only then can it be counted. Counting *wou* without a classifier is ungrammatical, as shown in (7c).

- (7) a. *wou*                      *quu'*                      *raq*  
          mango-hood              FRUIT:ROUND              two  
          ‘two mango fruits’
- b. \* *wou*  
      mango-hood
- c. \* *wou*                      *raq*  
      mango-hood      two

Long-shaped fruits like bananas take the fruit classifiers *kam* ‘CLF.FRUIT:LONG’ and tubular fruits take *yis* ‘CLF.FRUIT:CYLINDRICAL’. *Kam* and *yis* have the same referential and distributional properties as *quu*, as illustrated in (8) and (9).

- (8) a. *muxui kam 'ut*  
 banana-hood FRUIT:LONG four  
 ‘four bananas’
- b. \* *muxui*  
 banana-hood
- c. \* *muxui ut*  
 banana-hood four

In (9a), the classifier *yis* combines with a nominal compound *tei qar*, composed of the generic noun *tei* ‘tree’ and *qar* ‘food’, which expresses everything related to cassava (plant, leaf, tuber, etc.). Again, the nominal *tei qar* on its own is not a referential expression, (9b); rather, it must combine with the individuating classifier *yis*, compare (9a) and (9c). In this respect, the compound *tei qar* behaves just like *wou* in (7) and *muxui* in (8).

- (9) a. *tei qar yis yusan*  
 tree food FRUIT:CYLINDRICAL five  
 ‘five cassava tubers’
- b. \* *tei qar*  
 tree food
- c. \* *tei qar yusan*  
 tree food five

Fruit classifiers are also obligatory in questions about quantities, as shown in (10).

- (10) a.    *wat*            *quu'*            *yiran?*  
          coconut-hood FRUIT:ROUND how.many/much  
          ‘How many coconuts?’
- b. \*    *wat*                            *yiran?*  
          coconut-hood                how.many/much

In short, plant names are non-referential and must combine with another lexeme to become referential and countable. In the preceding examples, the intended referent was a fruit, and an obligatory fruit classifier was used. If the intended referent is some other part of the plant, the plant name combines with a noun that expresses that part of the plant’s whole, including *bag* ‘seed’, *wa* ‘leaf’, *qaau* ‘flower’, or *heer* ‘stem’. The part-of-whole noun *bag* ‘seed’ also functions as a numeral classifier, and is discussed in the next section. The other part-of-whole nouns have grammatical properties that differ from numeral classifiers. This is further discussed in section 2.

### **1.3. The general classifier *bag* ‘CLF’ < ‘seed’**

Teiwa has one ‘general’ numeral classifier: *bag*, which originally means ‘seed’. As a general classifier *bag* functions to classify nouns that are outside the semantic domains of the other, semantically more specific, shape classifiers for fruits and tubers, and the human classifier. For example, *bag* can optionally classify nouns that do not have a dedicated classifier, such as animals or non-plant objects. Zubin and Shimojo (1993) refer to this function as the ‘complement’ function of general classifiers. In addition, *bag* can also substitute for other, semantically more precise classifiers. For example, a speaker can classify mango fruits with *bag* instead of the dedicated fruit classifier *quu'*. This function is referred to as the ‘default’ function of a general classifier by Zubin and Shimojo (1993). While the Teiwa general classifier *bag* is most often used in the complement function, it can also be used in the default function.

The use of *bag* as general classifier is apparent in natural discourse and texts. In a corpus of about one hour of various narratives and conversations (Klamer 2010:34), *bag* is

used in its original part-of-whole sense of ‘seed’, but also to count children, eyes, and fish. However, numeral expressions in the corpus are quite rare, and they cover a limited range of semantic domains, so that an additional data set was collected through a field experiment. The experiment was designed to elicit numeral expressions with a semantically wide range of nouns and is described below. The results indicate that (i) *bag* is grammatically optional; (ii) the semantics of *bag* ‘seed’ are bleached; and (iii) the use of *bag* varies across individuals.

#### *Description of the experiment*

The experiment took place in two stages: one pilot experiment in 2010 (by Laura Robinson), and a second one in 2011 (by the author). Both experiments were carried out in the Teiwa speaking village of Lebang, on Pantar island. In both experiments, adult native speakers of Teiwa were asked to describe 43 color picture stimuli designed by Antoinette Schapper (2010b) to elicit classifier constructions in Alor Pantar languages. The 43 pictures depict various objects (natural, artificial, modern, traditional), animals (mammals, insects) and humans, in various types of configurations (standing, lying, hanging, flying, etc.), focusing on the lower numbers. Some of the pictures are illustrated below. In the experiment, the linguist and the speaker both sat facing a laptop screen, on which the pictures were displaced one by one. When a picture was shown, the linguist invited the speaker to describe the content of the picture by asking *Berapa (ada)?* ‘How many (are there)?’. The sessions were video recorded, and all responses were transcribed, translated, glossed, and entered into Toolbox.

In the first run of the experiment, 2 speakers participated. The results showed much inter-speaker variation in the use of classifiers in numeral expressions. A year later, the same experiment was run with 4 more speakers, who showed very similar inter-speaker variation, and the results of both experiments were combined into one data set containing  $6 \times 43 = 258$  numeral expressions. A few times, speakers offered a second, alternative construction at the time of recording, and these constructions were also included in the data set.

### *Summary of the results*

Summaries of the results are presented in Table I-IX below. They lead to the following observations. Firstly, in the experimental context, *bag* is the most frequently used classifier in Teiwa, and it is grammatically optional: none of the pictured objects has numeral expression where *bag* is used in 100% of the utterances (Table I-IX). An illustration of the optionality of *bag* is given in (11)-(12), which are the responses two speakers gave to picture 34 of the stimuli set (Figure 3).<sup>6</sup>



Figure 3. Picture 34 of the stimuli set: 4 arrows.

(11) *qafilat bag 'ut pial ga'*  
arrow CLF four stick along  
'four arrow (blades) with sticks' (MT)

(12) *qafilat 'ut*  
arrow four  
'four arrow (blades)' (OT)

Secondly, the experimental results suggest that *bag* is indeed derived from the part-whole noun 'seed': it is typically used to classify peanuts, grains of corn, grains of rice and tamarind seeds (Table I), but not fruits (Table II). However, one speaker classified the mango fruits with *bag*, thereby substituting it for the dedicated fruit classifier *quu* 'CLF.FRUIT:ROUND'. This is an instance of the 'general' classifier use of *bag* noted above.

---

<sup>6</sup> In the examples, capitals in brackets indicate speaker's initials.

Table I. Seeds, grains and fruits counted with *bag*.

<i>Objects on picture</i>	<i>Expression</i>	<i>No/speakers using expression with bag</i>	<i>% (N = 6)</i>
unpeeled peanuts	<i>moxo' qas bag</i> earth bean CLF	5	83
corn seeds	<i>batar bag</i> corn CLF	3	50
uncooked rice seeds	<i>raax/qar dig bag</i> rice/food peeled.seed CLF	2	33
tamarind seeds (bunch of 2, hanging in tree)	<i>tami bag</i> tamarind CLF	1	17
mango fruits (on table)	<i>wou bag</i> mango CLF	1	17

Table II. Fruits and bunches of fruit counted without *bag*

<i>Objects on picture</i>	<i>Expression</i>	<i>No/speakers using expression with bag</i>	<i>% (N = 6)</i>
corn cobs (peeled, on floor)	<i>batar kir</i> corn ear	0	0
mango fruits (in tree)	<i>wou quu'</i> mango FRUIT:ROUND	0	0
grape bunches	<i>tei yis baai/kiil</i> wood fruit bunch/stalk	0	0
banana bunches (in tree)	<i>muxui pluxun</i> banana hand.of	0	0
banana fruits (on table)	<i>muxui kam/yis</i> banana FRUIT:LONG/ FRUIT:CYLINDER	0	0
coconut bunches (in tree)	<i>wat baai</i> coconut bunch	0	0
coconuts (without skin, lying on floor)	<i>wat quu'</i> COCONUT FRUIT:ROUND	0	0

Tables III through IV below show that *bag* is also used to classify artifacts such as arrows, necklaces, windows, or chairs. In addition, it can classify animals like mice, buffaloes, bats and snakes, as in Table VIII. In these contexts, *bag* classifies nouns outside

the semantic domains of the other Teiwa classifiers: it has lost its ‘seed’ meaning completely and functions as a general numeral classifier. Apart from having no lexical meaning in these contexts, it is also optional. Speakers count the artifacts in Table IV with *bag*, but do not use *bag* when counting the artifacts in Table V. Also, pointed natural objects like sugarcane or bamboo sticks are counted with *bag*, as in Table VI, but *bag* is not used to count wooden sticks, trees or planks, as in Table VII. Similarly, the animals in Table VIII occur with *bag*, but those in Table IX do not. In short, *bag* is used as an optional general classifier here.

Table III. Dowry/bride prize items counted with *bag*

<i>Objects on picture</i>	<i>Expression</i>	<i>No/speakers using expression with bag</i>	<i>% (N = 6)</i>
arrows side by side on table	<i>qafilat bag</i> arrow CLF	5	83
bracelets on table	<i>Gelang/kenat bag</i> bracelet(IND)/bracelet CLF	4	67
bells (3 sets) on table	<i>gilan bag</i> bell CLF	3	50
necklaces on table	<i>bof (tar) bag</i> necklace (rope) CLF	2	33
bronze drums	<i>belian/quan bag</i> k.o. bronze drum CLF	2	33

Table IV. Other artifacts counted with *bag*

<i>Objects on picture</i>	<i>Expression</i>	<i>No/speakers using expressions with bag</i>	<i>% (N = 6)</i>
windows in wall	<i>Jendela/tua' bag</i> window(IND)/window CLF	5	83
chairs in room	<i>kadera bag</i>	3	50
lime container	<i>kapar bag</i>	3	50
plastic water bottle	<i>botol bag</i>	1	17
woven basket	<i>koop/tanaat/soqat bag</i> k.o.woven basket CLF	1	17

Table V. Artifacts NOT counted with *bag*

<i>Objects on picture</i>	<i>Expression</i>	<i>No/speakers using expression with bag</i>	<i>% (N = 6)</i>
houses in village	<i>yaf heer</i> house base	0	0
folded pieces of cloth	<i>kian/ nuan (qap)</i> k.o.cloth cut	0	0
wound plastic ropes	<i>tar yeni/ikat (IND)</i> rope wound/wound(IND)	0	0

Table VI. Wooden/natural objects with long/pointed shape counted with *bag*

<i>Objects on picture</i>	<i>Expression</i>	<i>No/speakers using expression with bag</i>	<i>% (N= 6)</i>
sugarcane sticks	<i>wux bag</i>	3	50
bamboo sticks	<i>yaa' bag</i>	1	17

Table VII. Wooden/natural objects with long/pointed shape NOT counted with *bag*

<i>Objects on picture</i>	<i>Expression</i>	<i>No. of speakers using bag</i>	<i>% (N = 6)</i>
trees in a field	<i>tei heer</i> tree base	0	0
planks flat on ground	<i>hala'tei qap</i> plank/tree cut <i>papan qap</i> plank(IND) cut	0	0
wooden sticks	<i>tei baq</i> tree stick	0	0

Table VIII. Animals counted with *bag*

<i>Animals on picture</i>	<i>Expression</i>	<i>No. of speakers using bag</i>	<i>% (N = 6)</i>
mice	<i>dur bag</i>	3	50
water buffaloes	<i>qarbau bag</i>	3	50
insect	<i>reed/dalidi/ta'an/or bag</i> butterfly/wasp/bee CLF	2	33
bats (hanging)	<i>madi bag</i>	2	33
bats (flying)	<i>madi bag</i>	1	17
small animals (in tree)	<i>kamau/mauqubar/urax/dur/rau bag</i> wild.cat/frog/snail/mouse/civet.cat CLF	1	17
snakes	<i>daam bag</i>	1	17



Table IX. Animals NOT counted with *bag*

<i>Animals on picture</i>	<i>Expression</i>	<i>No. of speakers using bag</i>	<i>% (N = 6)</i>
wild cat	<i>kamau</i>	0	0
frog	<i>mauqubar</i>	0	0
snail	<i>urax</i>	0	0
civet cat	<i>rau</i>	0	0
insect	<i>ta'an/reed</i> wasp/butterfly	0	0

In Table III, I listed the items that I expected to typically occur in numeral expressions because they are part of the traditional Teiwa bride prize and dowry negotiations: arrows, bracelets, ankle bells, bronze drums and necklaces. However, in the two spontaneous narrative texts about bride prize and dowry that I collected in 2011, none of these objects occurred with a classifier, which indicates that even in frequently used and conventionalized numeral constructions *bag* is grammatically optional.

Finally, the use of *bag* as numeral classifier varies across individuals, see Table X. There are indications that speakers over 50 use *bag* more sparingly (5-11% of the utterances) than speakers under 40 (25-42% of the utterances). However, the number of speakers is too small to make inferences about differences between age groups. Note also that the patterns of the speaker born in 1989 are similar to those of the older generation, so that, if any trend towards increased use of *bag* exists in younger speakers, it is weak at most. What the data clearly show, however, is that there is significant inter-speaker variation in the use of *bag*. *Bag* is not only grammatically optional, its use also varies per speaker.

Table X. The use of *bag* per speaker, ordered by year of birth

<i>Year of birth</i>	<i>1948</i>	<i>1960</i>	<i>1972</i>	<i>1981</i>	<i>1988</i>	<i>1989</i>
Number of objects counted with <i>bag</i> (N=43)	5	2	17	18	12	7
<b>Percentage</b>	11,6	4,7	39,5	41,9	25,5	16,3

Speakers under 40 use *bag* with a wider range of referents than the speakers over 50. A pairwise comparison of the utterances of the speakers showed that overall, the younger

speakers tend to use *bag* where older speakers do not use a classifier. This is illustrated with the responses to picture 40 (Figure 4) in (13) and 38 (Figure 5) in (14).



Figure 4. Picture 40 of the stimuli set: 6 grains of rice.

(13) a. *raax dig tiaam*  
rice peeled.seed six  
'six [grains of] uncooked rice' (OT, 1960)

b. *raax dig bag tiaam*  
rice peeled.seed CLF six  
'six grains of uncooked rice' (LL, 1989)



Figure 5. Picture 38 of the stimuli set: 4 water buffaloes.

- (14) a. *qarbau*            *ut*  
 water.buffalo    four  
 ‘four water buffaloes’ (OT, 1960)
- b. *Qarbau*                            *ut*    *ur*    *gom*    *ma*    *tewar.*  
 water.buffalo                            four    field    inside    come    walk  
 ‘Four water buffalos walk in the field’ (MT, 1948)
- c. *qarbau*            *bag*    *ut*  
 water.buffalo    CLF    four  
 ‘four water buffaloes’ (LT, 1981)
- d. *qarbau*                            *bag*    *ut*  
 water.buffalo                            CLF    four  
 ‘four water buffaloes’ (JMW, 1988)

In sum, the general classifier *bag* is optional and can be used to classify everything except fruits; including humans (in particular children), animals, and non-plant objects. The ‘seed’ semantics of the part-of-whole noun *bag* ‘seed’ from which it developed have been bleached, perhaps more so for the speakers under 40 than for those over 50 years of age.

#### **1.4. Summary**

Teiwa has a small set of five classifiers, of three types: one to classify humans, three to classify fruits according to their shape, and one general classifier. The human classifier is obligatory in questions about quantities and in answers to such questions, and is otherwise optional. Fruit classifiers are always obligatory, and the general classifier is always optional. The general classifier derives from the noun ‘seed’ but when it combines with nouns referring to children, animals and inanimate objects, it has lost its ‘seed’ semantics.

## 2. The development of Teiwa classifiers

This section addresses the question: Where did the Teiwa classifiers originate from? I argue that (i) they were not inherited from the ancestor language proto-Alor Pantar, and (ii) that (except for the human classifier *-man*) they developed out of a particular class of nouns, the part-of-whole nouns, which was inherited. Both arguments are based on a comparison of the Teiwa classifiers with the lexicon of languages that are members of the same language family, the Alor Pantar family (Holton et.al. 2012).

Numeral classifiers are found throughout the Alor Pantar family. Apart from Teiwa, we find classifiers in its relatives across the two islands, including: Western Pantar (15), Adang (16), Klön (17), Abui (18) and Kamang (19). A human classifier is found in some but not all Alor Pantar languages (Teiwa *-man*, Klön *ana*, Abui *ning*, Kamang *ning*). None of the Alor Pantar languages has a dedicated classifier for animals.<sup>7</sup> Of the Alor Pantar languages, Adang has the richest set of classifiers (16), while Western Pantar (15) is unique in recruiting classifiers from nominal as well as non-nominal lexemes.

(15) Western Pantar (Holton, to appear)

Form	Meaning	Classifies
<i>bina</i>	‘detached’	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

(16) Adang (Robinson and Haan, to appear)

Form	Meaning	Classifies	Combines with
<i>pa’</i>	‘non-round fruit’	general classifier for objects of many shapes and sizes	arrows, drums, borrowed nouns, birds, fish <sup>8</sup>

<sup>7</sup> If animals are classified this is done with a classifier that is also used to classify fruits and humans (Adang *pir*) or inanimate objects (Abui *lohu*), see the data below.

<sup>8</sup> 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).

<i>beh</i>	leaf	flat, flexible objects	leaves, money notes
<i>bo'</i>	none	flat, rigid, large objects	fields
<i>bo'</i>	log	long, cylindrical, rigid objects	bamboo, logs
<i>'ahang</i>	slice	flat, rigid, small objects	wood, walls
<i>'anemeng</i>	sheet	non-flat, flexible objects	clothes, rope, string
<i>el</i>	(none)	rigid, standing objects	buildings, trees
<i>'afail</i>	seed	small rigid objects	corn kernel, rice grain
<i>'ir</i>	(none)	long, cylindrical, jointed objects	bamboo, sugarcane
<i>kumang</i>	piece	short, cylindrical objects	logs, eels, snakes
<i>pir</i>	round fruit	round objects	fruits, animals, people
<i>puh</i>	(none)	hanging objects	banana blossoms, corn ears
<i>tuling</i>	drop <sup>9</sup>	liquids	water, oil, wine
<i>bar</i>	bunch	short, clustered, hanging objects	coconut bunches, earrings, bells, betel nut

(17) Klon (Baird 2008)

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

(18) Abui (Schapper 2010a)

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

(19) Kamang (Schapper, to appear)

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

The data in (15)-19) allow us to conclude that (i) numeral classifiers are common in Alor Pantar languages, but (ii) they have variable shapes and make different semantic classifications, so that (iii) no classifier is reconstructable for proto-Alor Pantar.

On the other hand, we do find cognate lexemes in the class of part-of-whole nouns. A cursory comparison of the Teiwa part-of-whole nouns with similar nouns in other Alor-Pantar languages rendered the cognates in (20), which exhibit regular sound

<sup>9</sup> This is a mensural numeral quantifier.

correspondences (cf. Holton et. al. 2012). In addition, irregular forms like those in (21) are also attested.<sup>10</sup>

(20)	<b>Teiwa</b>	<b>Kaera</b>	<b>W Pantar</b>	<b>Adang</b>	<b>Klon</b>
a.	<i>heer</i> ‘stem, base’	<i>er</i> ‘stem, base’	<i>haila</i> ‘base, area’	<i>el</i> ‘rigid, standing object’	<i>yar</i> ‘trunk’
b.	<i>kiil</i> ‘stalk’ (long, thin, flexible’			<i>ir</i> ‘articulation of bamboo, CLF for rigid, long, jointed objects’	
(21)	<b>Teiwa</b>	<b>Kaera</b>	<b>W Pantar</b>	<b>Adang</b>	<b>Klon</b>
a.	<i>wa</i> ‘leaf’	<i>wa</i> ‘leaf’	<i>waya</i> ‘leaf’	<i>beh</i> ‘leaf’	
b.	<i>bag</i> ‘seed’, CLF			<i>pa</i> ‘non-round fruit’; CLF	
c.	<i>qap</i> ‘a cut’ (soft/hard material)		<i>kakka</i> ‘long, stiff, flat object’		<i>kak</i> ‘board, plank’

From the comparative evidence we can make the following inferences relating to the Teiwa classifiers. First, as no numeral classifiers can be reconstructed to proto-Alor Pantar, we have no evidence to assume that Teiwa classifiers are inherited. However, there is evidence that the language ancestral to Teiwa, Kaera, Western Pantar, Adang and Klon had a set of part-of-whole nouns, including those with the meanings in (20)-(21). In Teiwa, at least one of the inherited part-of-whole nouns, *bag*, has developed into a numeral classifier.

This diachronic perspective is supported by synchronic evidence that part-of-whole nouns and fruit classifiers in Teiwa share a similar function: both create referential nominal expressions and individuate entities. However, the distribution of the part-of-whole nouns is less restricted than the fruit classifiers, which can only appear in numeral contexts. The evidence is presented in what follows.

---

<sup>10</sup> Note that cognate part-of-whole nouns are found among the languages of Pantar and western Alor; the central-eastern languages Abui and Kamang use different forms.

As mentioned in section 1.2, Teiwa plant names are non-referential and must combine with either a fruit classifier, or with another noun -- if the referent is not the fruit but some other part of the plant. An illustration is (22), where the generic plant name *wou* (glossed as ‘mango-hood’, cf. section 1.2) combines with the part-of-whole nouns *bag* ‘seed’, *wa* ‘leaf’, *qaau* ‘flower’, and *heer* ‘stem, base’.

(22)	<i>wou</i>	<i>bag</i>	<i>wou</i>	<i>wa</i>
	mango-hood	seed	mango-hood	leaf
	‘mango seed(s)’		‘mango leaf(s)’	
	<i>wou</i>	<i>qaau</i>	<i>wou</i>	<i>heer</i>
	manggo-hood	flower	manggo-hood	stem
	‘mango flower(s)’		‘mango tree(s)’	

Other parts-of-whole nouns are *kiil* ‘stalk’, *baq* ‘stick’, and *qap* ‘cut’, (23)-(24). A list of frequently used part-of-whole nouns in Teiwa is given in Table 5.

(23)	<i>tei</i>	<i>kiil,</i>	<i>tei</i>	<i>baq,</i>	<i>tei</i>	<i>heer</i>
	tree-hood	stalk	tree-hood	stick	tree-hood	stem
	‘twig(s)’		‘wooden stick(s)’		‘tree trunk(s)/stem(s)’	
(24)	<i>wurax</i>	<i>qap,</i>	<i>wurax</i>	<i>baq</i>		
	rattan-hood	cut	rattan-hood	stick		
	‘piece(s) of split rattan’		‘rattan stalk(s)’			

Table 5. Teiwa part-of-whole nouns with their meaning and the object they refer to

<b>Form</b>	<b>Meaning</b>	<b>Refers to</b>
<i>wa'</i>	leaf	leaves, leaf-shaped objects
<i>heer</i>	stem	stem, base, objects with areal extent
<i>qaau</i>	flower	flowers, flower-shaped objects
<i>kiil</i>	stalk	long, thin and flexible objects (e.g. rattan)
<i>baq</i>	stick	long, stiff objects (e.g. sticks, trunks)
<i>qul</i>	top	top of plant or tree
<i>qap</i>	cut	flat objects (hard, e.g. wooden plank, or soft, e.g. cloth)
<i>bag</i>	seed	seeds
<i>dig</i>	seed (peeled)	peeled seeds (e.g. uncooked rice)
<i>kir</i>	ear (of corn)	corn cob with leaves peeled off

While both part-of-whole nouns and fruit classifiers function to create referential nominal expressions, fruit classifiers can only appear in numeral contexts, while part-of-whole nouns are less restricted; compare ungrammatical (25a) and grammatical (26a), (27a).

- (25) a.\*    *wou*            *quu'*  
                  mango-hood    FRUIT:ROUND
- b.        *wou*            *quu'*            *nuk*  
                  mango-hood    FRUIT:ROUND    one  
                  'a / one mango'
- c.        *wou*            *quu'*                            *yusan*  
                  mango-hood    FRUIT:ROUND    five  
                  'five mangos'
- (26) a.        *wou*            *wa'*  
                  manggo-hood    leaf  
                  'mango leaf(s)'
- b.        *wou*            *wa'*            *yusan*  
                  manggo-hood    leaf            five  
                  'five mango leaves'



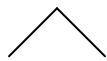
- (27) a. *batar bag*  
 corn seed  
 ‘corn seed(s)’
- b. *batar bag nuk*  
 corn seed one  
 ‘a / one corn seed’

Fruit classifiers and part-of-whole nouns also occupy different positions in the Teiwa NP. A Teiwa NP is maximally composed of an initial head noun ( $N_{\text{HEAD}}$ ), followed by an attribute (ATTR) (e.g. an adjective), numeral (NUM) expressions of quantity, a demonstrative (DEM) and a demonstrative particle (PART), see (28a).<sup>11</sup> A classifier and numeral form a separate numeral phrase, (28b). In contrast, a part-of-whole noun is part of the compound noun that heads the nominal phrase, (28c). The construction in (28b) is illustrated in (29), the one in (28c) is illustrated in (30).

(28) a. [ $N_{\text{HEAD}}$  ATTR NUM DEM PART]<sub>NP</sub>

b. [ $N_{\text{HEAD}}$  ATTR [CLF NUM] NUMP DEM PART]<sub>NP</sub>

c. [ $N_{\text{HEAD}}$  ATTR NUM DEM PART]<sub>NP</sub>



N N

plant part-of-whole

<sup>11</sup> In the demonstrative position, we often find *ga'an* (glossed as ‘that.KNWN’), a 3SG object pronoun that also functions as a demonstrative modifier of nouns. In the Particle slot are the demonstrative particles *u* ‘DISTAL’ and *a* ‘PROXIMATE’. These particles occupy the NP-final position, and they mark definiteness and/or the location of NP referent with respect to the speaker.

(29) [N<sub>HEAD</sub> ATTR [CLF NUM] DEM PART]<sub>NP</sub>  
*muxui muban kam yesraq ga'an u*  
 banana-hood ripe CLF.FRUIT:LONG seven that.KNWN DISTAL  
 'those seven ripe bananas...'

(30) [N<sub>HEAD</sub> ATTR NUM DEM PART]<sub>NP</sub>  
*wurax qap ii yerig ga'an u*  
 rattan cut red three that.KNWN DISTAL  
 'those three red pieces of rattan'

Apart from their different syntactic properties, Teiwa fruit classifiers and part-of-whole nouns also have different categorizing functions. Fruit classifiers can only be used to count fruits, while part-of-whole nouns are also used to count objects that are not plant-related. Examples are *heer* 'stem, base' in (31)-(33), *wa'* 'leaf' in (34)-(35) and *bag* 'seed' in (36)-(38).

*Heer* refers to the stem or base of a tree, but is also used to count bronze drums, houses and areas like gardens or rice fields:

(31) *belian heer nuk*  
 bronze.drum stem/base one  
 'one moko drum'

(32) *yaf heer yerig*  
 house stem/base three  
 'three houses'

(33) *maxar heer yerig*  
 new.garden stem/base three  
 'three new garden fields'

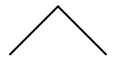
*Wa* ‘leaf’ is metaphorically extended to count entities that are leaf-like in shape, such as pieces of paper and money notes:

- (34) *kertas*                      *wa*’    *yerig*  
 paper (IND)                      leaf    three  
 ‘three sheets of paper’
- (35) *sen*                                      *wa*’    *yerig*  
 money (Malay<Dutch)              leaf    three  
 ‘three notes of money’

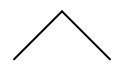
*Bag* ‘seed’ is used to count small seeds or seed-like entities but also for larger objects of a various kinds and shapes, including *qafilat* ‘arrow’ in (36) and *kadera* ‘chair’ in (37), as well as animals like *qarbau* ‘water buffalo’ in (38).

- (36) *qafilat bag*    *’ut*  
 arrow seed    four  
 ‘four arrows’
- (37) *kadera*                      *bag*    *yesraq*  
 chair (Portuguese)    seed    seven  
 ‘seven chairs’
- (38) *qarbau*            *bag*    *’ut*  
 water.buffalo seed    four  
 ‘four water buffaloes’

Although differing in their synchronic properties, classifiers and part-of-whole nouns can occur in constructions that are identical at the surface. Compare for example (39), which has a part-of-whole noun *baq* ‘stick’, with (41), which has a classifier *bag* ‘seed’. Part-of-whole nouns could develop into classifiers through reanalysis of their structural position: from being part of a compound, as in (39), to being part of a numeral phrase, as in (41), via ambiguous structures like the one in (40), where *bag* may be analyzed as either a part-of-whole noun, or a numeral classifier.

(39) [ N<sub>HEAD</sub>                      NUM ]<sub>NP</sub>  
  
*tei*    ***baq***                      *yerig*  
 tree    stick                      three  
 ‘three sticks’

(40) [*batar* ***bag***]<sub>N</sub> *yerig*  
*batar* [***bag*** *yerig*]<sub>NUMP</sub>  
 corn    seed    three  
 ‘three corn seeds’

(41) [N<sub>HEAD</sub>                      NUMP ]<sub>NP</sub>  
  
*bala'*                      ***bag***    *yerig*  
 bed                      CLF    three  
 ‘three beds’

In sum, Teiwa part-of-whole nouns and numeral classifiers share the function of creating referential expressions, but the numeral classifiers have developed into more functional items through reanalysis of their phrasal position. Teiwa did not inherit its classifiers from an ancestor language, though it did inherit a set of part-of-whole nouns, at least one of which has been recruited to become a classifier. This account explains the origin of the general classifier *bag*, and I hypothesize that the fruit classifiers developed in a similar way.<sup>12</sup> The next section discusses a semantic motivation for this development.

---

<sup>12</sup> As we lack information about the lexical content of human classifier *-man* its origin remains unknown.

### 3. Classifiers and number neutral nouns

It is often suggested that languages with classifiers have nouns that are ‘number neutral’; that is, they have no morphology marking number on nouns. This classic observation goes back to Greenberg, who observed that languages which make use of numeral classifiers in their ‘basic mode of forming quantitative expressions’ never have compulsory number marking on the noun (1972/77: 286). An example of a language showing such a correlation is Indonesian, which has both classifiers and bare NPs that are number neutral. For example, the NP *anak* ‘child’ can be interpreted as a mass noun ‘child’, as a singular count noun ‘a child’, or as plural ‘children’. In ‘number neutral’ languages like Indonesian, an NP consisting of just a bare noun will be semantically neutral for number, and can be understood as either a mass or a count noun, and as either singular or plural (Gil 1987, 2011).

Correlating the notion of number neutrality with the use of classifiers, it is often observed that languages with number neutral nouns develop a category of classifiers to individuate the noun, and thus provide the necessary units for quantification of the noun (cf. Thompson 1965, Link 1991, Gil 2011). In this view, languages with number neutral nouns “need” classifiers for a semantic reason: to individuate nouns before they can be counted.

In light of this theory, it is relevant to mention that Teiwa (and most other Alor Pantar languages), are number neutral: nominal plurality is not indicated by morphology on the noun, but via a separate plural number word. Examples of plural number words in Alor Pantar languages are given in Table 6. Illustrations are given in (42)-(44).

Table 6. Plural number words in Alor Pantar (Klamer, Schapper & Corbett, to appear).

<b>Language</b>	<b>Form</b>	<b>Source</b>
Western Pantar	<i>maru(ng)</i>	Holton, to appear
Teiwa	<i>non</i>	Klamer 2010, Schapper & Klamer 2011
Klon	<i>onon</i>	Baird 2008
Abui	<i>loku</i>	Kratochvíl 2007, p.c.
Kula	<i>du(wa)</i>	Kratochvíl p.c.
Sawila	<i>do</i>	Kratochvíl p.c.
Kamang	<i>nung</i>	Schapper & Klamer 2011

Without a plural number word, the noun has a singular or plural interpretation; with a plural number word it can only be plural, compare (42a-b). The plural number word enforces a plural interpretation, (43)-(44).

Teiwa (Schapper & Klamer 2011)

- (42) a. *Qavif ita'a ma gi?*  
goat where come go  
'Where did (the) goat / goats go?'
- b. *Qavif non ita'a ma gi?*  
goat PL where come go  
'Where did (the) (several) goats go?'

Klon (Baird n.d.)

- (43) *Ge-ebeng non go-tohok.*  
3.GEN-friend PL 3-meet  
'(He) met his friends', \*'(He) met his friend'

Western Pantar (Holton, to appear)

- (44) *Raya marung wang hundar*  
chief PL exist amazed  
'The chiefs were amazed'

In sum, as a language with number neutral nouns, Teiwa nouns must be individuated before they can be counted, and developing a set of numeral classifiers helps to serve this need.

## 4. Teiwa classifiers in their areal context

In the previous sections I argued for a scenario where certain members of the Teiwa part-of-whole nouns developed into numeral classifiers through reanalysis of their phrasal position, a process that may have been enhanced by the number neutral status of the language.

This language internal development was probably reinforced by intensive contact with languages that have classifiers. One such language is Indonesian, the national language of Indonesia. Indonesian is the language of media and education, and is spoken as second language by virtually everyone on Pantar and Alor, while it 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.

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).<sup>13</sup> In this respect, Indonesian *buah* is thus quite similar to the general Teiwa classifier *bag*. Recent intensive contact with Indonesian could have spiraled the part-of-whole noun *bag* 'seed' into becoming a general classifier. Note, however, the Indonesian noun *buah* means 'fruit', and as a classifier it classifies objects and fruits, but no animals. In contrast, Teiwa *bag* originally means 'seed', and classifies objects and animals, but no fruits. The only feature shared by *buah* and *bag* is their general classifying function; this is the only part of Teiwa *bag* that has been copied from Indonesian.

Apart from recent contact with Indonesian, it is likely that the development of classifiers in Teiwa was also influenced by contact with other classifier languages. Obvious candidates are the Austronesian languages spoken in the vicinity of Alor and Pantar. Many

---

<sup>13</sup> *Buah* is a reflex of the proto-Malayo-Polynesian noun \**buaq* 'fruit', and cognate with the proto-Oceanic general classifier \**puaq*.

(if not most) Austronesian languages in eastern Indonesia have numeral classifiers,<sup>14</sup> examples include those in Table 7. 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, to appear).

Table 7. Austronesian languages with numeral classifiers in eastern Indonesia

<b>Region/island</b>	<b>Language</b>	<b>Source</b>
Flores	Rongga	Arka 2008
	Keo	Baird 2001, 2002
	Sika	Arndt 1931
Timor	Tetun Fehan	Williams-van Klinken 1999
	Tetun Dili	Williams-Van Klinken et. al. 2002
	Waimaha	Hull 2002
	Leti	Van Engelenhoven 2004 <sup>15</sup>
Aru	Dobel	Hughes 2000
	Kei	Geurtjens 1921
C Moluccas	Larike	Laidig & Laidig 1995
Halmahera	Taba	Bowden 2001
Sumba	Kambera	Klamer 1998

Beyond the languages of Indonesia in Table 7, classifiers are attested throughout the Austronesian-speaking world outside of Taiwan.<sup>16</sup> Examples include the Vietnam language

<sup>14</sup>Blust (2009: 282) remarks that '[classifiers] are reported sporadically in eastern Indonesia'. This is clearly an understatement, as virtually all the Austronesian languages in eastern Indonesia we know about either have independent numeral classifier lexemes, or numeral prefixes that are synchronically meaningless but clearly derive from classifiers (cf. Leti, as explained in the next footnote). It is beyond the scope of this article to go into details here, but the list in Table 7 would suggest that classifier languages are not 'sporadic' in eastern Indonesia.

<sup>15</sup> Although Van Engelenhoven does not mention numeral classifiers in his grammar, the noun *vua/vua* 'fruit' (2004: 447), a reflex of Proto-Austronesian *\*buaq*, must have assimilated with the numerals 2-9 as a (synchronically meaningless) prefix *v(o)-* (2004:165). Similar constructions 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).



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). 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 (45) (<Malayo-Polynesian \*buaq). Other reconstructed forms are the classifier for animate beings \*mwane, a classifier for wooden or elongated objects (\*kaiu) and one for persons (\*tau) (Lynch, Ross and Crowley 2002:74). The Oceanic languages with classifiers include those listed in (47) (Lynch et.al. 2002:73-74). Major subgrouping information is included in brackets.

(45) (Proto-Oceanic; Lynch et.al. 2002:74)

*ta puaq tolu a niuR*  
 ART fruit three ART coconut  
 ‘two coconuts’

- (46) 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. They must have been quite commonly used in proto-Oceanic. For instance, Lynch et. al mention that a word like \*niuR ‘coconut’

---

<sup>16</sup> 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 below, it refers to “the Malayo-Polynesian subgroup of Austronesian”.

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 (2002:74).

In short, 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. This makes it likely that contact with Austronesian languages has contributed to the development of numeral classifiers in the Alor Pantar languages listed in (15)-(19) above. The contact did not involve diffusion of lexemes: no similarity in shape or semantics exists between classifiers in Alor Pantar languages and classifiers of Austronesian languages in the area (compare the difference between Indonesian *buah* and Teiwa *bag*). Yet, it is remarkable that the part-of-whole nouns/classifiers reconstructed for proto-Oceanic function in a way that is very much like how they are used in Alor Pantar languages such as Teiwa today. For example, Teiwa *wat* ‘coconut-hood’ is a generic noun just like proto-Oceanic \*niuR: both refer to ‘coconut-hood’ and must combine with a classifier to refer to the actual fruit.

Supporting evidence for the areal pressure from Austronesian into Papuan languages comes from the other Papuan families in eastern Indonesia with longstanding contacts with Austronesian.<sup>17</sup> First, on Timor, the Papuan languages Makasae and Makalero have classifiers coupled with many other traces of Austronesian influence (see Huber 2008, 2011). Second, many Papuan languages of the Bird’s Head have classifiers, including Abun (Berry & Berry 1999), Tehit (Flassy 1991), Maybrat (Dol 1999), Hatam (Reesink 1999). This is often coupled with Austronesian influence in word order, pronouns, numerals and lexicons (Voorhoeve 1989).<sup>18</sup> Third, in Halmahera, the West Papuan languages Tidore (Van Staden 2000) and Tobelo (Holton 2003) also have classifiers, and old Austronesian loans are found throughout the family, suggesting a long period of contact dating to the original settlement of the area by Papuan speakers (Voorhoeve 1994).

In short, in all the well-known zones of Austronesian-Papuan contact in eastern Indonesia numerous Papuan languages with classifiers are found. This is striking because classifiers are generally lacking in Papuan languages, and the few Papuan languages that do

---

<sup>17</sup> See the discussion in Klamer, Reesink & Van Staden 2008, Holton & Klamer, to appear and references cited there.

<sup>18</sup> Austronesian Biak (Van den Heuvel 2006) and other Austronesian languages of the Cendrawasih Bay were used as languages of wider communication before the advent of Indonesian.

have them occur in scattered locations throughout New Guinea (Aikhenvald 2000:123).<sup>19</sup> This suggests that classifiers are not a typical part of Papuan language structures, and it is not an accident that they are mostly found Papuan families with a long history of contact with Austronesian.

## 5. Summary and conclusions

Teiwa has five classifiers, of three types: one to classify humans, three to classify fruits according to their shape, and one general classifier. Fruits are the only objects that are obligatorily counted with classifiers. The human classifier is obligatory only in questions with the interrogative quantifier *yiran* ‘how many/much’, and in answers to such questions. The general classifier is always optional. Teiwa has no dedicated classifier for animals nor for inanimate objects; these can optionally occur with the general classifier *bag*. The general classifier derives from the noun ‘seed’. It can combine with nouns of a wide semantic range, including children, animals and inanimate objects.

A comparison of the Teiwa classifiers with the lexicon of closely related languages suggests that it is unlikely that Teiwa inherited this small and variable set of classifiers from its ancestor language. However, there is evidence that the ancestor of the Alor Pantar languages had a set of nouns referring to parts of plants (e.g. ‘stem’, ‘leaf’, ‘stalk’). Some of these nouns developed into numeral classifiers in Teiwa, through reanalysis of their position in the numeral NP. A semantic motivation supporting this development is that Teiwa has number neutral nouns which must be individuated before they can be counted, a function that is fulfilled by numeral classifiers.

The genesis of Teiwa classifiers must have been reinforced by contact with Austronesian languages that have classifiers. Ancient contact between the Alor Pantar languages and the Austronesian languages of eastern Indonesia must have enhanced the development of classifiers in the Alor Pantar languages, while recent and intensive contact

---

<sup>19</sup> Numeral classifiers are absent from the overviews of Papuan features by Foley (1986, 2000) and Aikhenvald & Stebbins (2007). Aikhenvald (2000:123) mentions ten Papuan languages with classifiers in scattered locations of Papua New Guinea: Iwam, Abau (East Sepik province), Chambri, Wogamusin, Chenapian (Lower Sepik), Angave, Tanae (Gulf Province), Folopa (Highlands), Wantoat, Awará (Morobe province).

with Indonesian may have spiralled one of them to become a general classifier like Indonesian *buah*.

By connecting the numeral classifier system of Teiwa with those of its sister languages of the Alor Pantar family and the wider linguistic context of eastern Indonesia, we can see that multiple factors are involved when classifiers are born in a language.

## References

- Aikhenvald, Alexandra Y. 2000. *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 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." *The Vanishing Voices of the Pacific Rim* ed. by O. Miyaoka, O. Sakiyama and M.E. Krauss, 239-66. Oxford: Oxford University Press.
- 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.
- Baird, Louise. 2002. *A grammar of Kéo: An Austronesian language of East Nusantara*. PhD dissertation Australian National University.
- Baird, Louise. 2008. *A Grammar of Klon: A Non-Austronesian Language of Alor, Indonesia*. Canberra: Pacific Linguistics.
- Baird n.d. Klon texts. Toolbox database, Leiden University.
- 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
- Chung, Siaw-Fong. 2010. "Numeral classifier *buah* in Malay: A corpus-based study." *Language and Linguistics* 11.3: 553-577.

- 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.
- Foley, William A. 2000. "The languages of New Guinea." *Annual Review of Anthropology* 357-404.
- Geurtjens, Hendrik. 1921. "Spraaikleer der Keieesche taal." *Verhandelingen van het Bataviaansch Genootschap voor Kunsten en Wetenschappen* 63.2. Albrecht & Martinus Nijhoff.
- Gil, David. 1987. "Definiteness, Noun-Phrase Configurationality, and the Count-Mass Distinction." In *The Representation of (In)definiteness* ed. by Erik .J. Reuland and Alice G.B. ter Meulen, 254-269. Cambridge, MA: MIT Press.
- Gil, David. 2011. "Numeral classifiers." In *The World Atlas of Language Structures Online* ed. by Matthew Dryer and Martin Haspelmath. München: Max Plank Digital Library. Available online at <http://wals.info/>. Accessed on 2013-02-11.
- Greenberg, Joseph. 1972 (1977). "Numeral classifiers and substantival number: problems in the genesis of a linguistic type." In *Stanford papers on language universals* 9, 1-39. Reprinted in A. Makkai et. al. (eds.), 1977. *Linguistics at the crossroads*, 276-300. Lake Bluff, Illinois, Padova: Jupiter Press.
- Haan, Johnson Welem. 2001. *A grammar of Adang: a Papuan language spoken on the Island of Alor, East Nusa Tenggara, Indonesia*. Ph.D. dissertation University of Sydney.
- van den Heuvel, Wilco. 2006. *Biak: A description of an Austronesian language of Papua*. PhD dissertation Vrije Universiteit Amsterdam. Utrecht: LOT dissertations.
- Himmelman, Nikolaus. 2005. "The Austronesian languages of Asia and Madagascar: Typological characteristics". *The Austronesian languages of Asia and Madagascar* ed. by Alexander Adelaar and Nikolaus P. Himmelman, 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. 2003. *Tobelo*. Munich: LINCOM Europa.
- Holton, Gary. To appear. "Sketch of Western Pantar (Lamma)". *Papuan Languages of Timor-Alor-Pantar: Sketch grammars* ed. by Antoinette Schapper. Canberra: Pacific Linguistics.
- 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 *Oceania* ed. by Bill Palmer. Berlin-New York: Mouton De Gruyter.
- Hopper, Paul J. 1986. "Some discourse functions of classifiers in Malay." In *Noun Classes and Categorization* ed. Colette G. Craig, 309-325. Amsterdam: John Benjamins.
- Huber, Juliette. 2008. *First steps towards a grammar of Makasae*. Munich: LINCOM Europa.
- Huber, Juliette. 2011. *A grammar of Makalero: A Papuan language of East Timor*. Ph.D. dissertation Leiden University.
- Hughes, Jock. 2000. "The Morphology of Dobel, Aru, with Special Reference to Reduplication." In *Spices from the East: Papers in Languages of Eastern Indonesia* ed. by Chuck E. Grimes, 131-180. Canberra: Pacific Linguistics.
- Hull, Geoffrey. 2002. *Waimaha*. Instituto Nacional de Linguística, Universidade Nacional de Timor Lorosa'e.
- Klamer, Marian. 1998. *A grammar of Kambara*. Berlin New York: Mouton de Gruyter.
- Klamer, Marian. 2010. *A grammar of Teiwa*. Berlin New York: Mouton de Gruyter.
- Klamer, Marian, Ger P. Reesink and Miriam van Staden. 2008. "East Nusantara as a linguistic area." In *From Linguistic Areas to Areal Linguistics*, ed. by Pieter Muysken, 95-150. Amsterdam: John Benjamins.

- Klamer, Marian, Antoinette Schapper and Greville Corbett. To appear. "Plural number words". In *Alor Pantar languages: History and Typology* ed. by Marian Klamer. Leiden: Brill.
- Kratochvíl, František. 2007. *A grammar of Abui', a Papuan language of Alor*. Ph.D. thesis Leiden University. Utrecht: LOT Publications.
- Laidig, Wyn D. and Carol J. Laidig. 1995. "A synopsis of Larike phonology and syntax." In *Descriptive Studies in Languages of Maluku* ed. Wyn D. Laidig, 19-42. NUSA 2.
- Link, Godehard. 1991. "Quantity and Number." In *Semantic universals and universal semantics* ed. D. Zaefferer, 133-149. Berlin: Foris Press.
- 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.
- Nishiyama, Kunio, and Herman Kelen. 2007. *A grammar of Lamaholot, Eastern Indonesia. The Morphology and Syntax of the Lewoingu Dialect*. München: Lincom Europa.
- Reesink, Ger P. 1999. *A grammar of Hatam, Irian Jaya, Indonesia*. Canberra: Pacific Linguistics.
- Robinson, Laura and John W. Haan. To appear. "Adang." *Papuan Languages of Timor-Alor-Pantar: Sketch grammars*, ed. by A. Schapper. Canberra: Pacific Linguistics.
- 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. 2010a. *Abui Numerals Questionnaire*. Unpublished ms. Leiden University.
- Schapper, Antoinette. 2010b. *Berapa? Numeral classifier photo stimuli*. Leiden University.
- Schapper, Antoinette. To appear. "Kamang." *Papuan Languages of Timor-Alor-Pantar: Sketch grammars*, ed. by A. Schapper. Canberra: Pacific Linguistics.
- 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. 2011. "Plural words in Papuan languages of Alor-Pantar." *Proceedings of Language Documentation & Linguistic Theory (LDLT)* 3, 247-256. London: SOAS.
- Schapper, Antoinette and Marian Klamer. To appear. "Numeral systems." In *Alor Pantar languages: History and Typology* ed. by Marian Klamer. Leiden: Brill.
- van Staden, Miriam. 2000. *Tidore: A Linguistic Description of a Language of the North Moluccas*. Ph.D. thesis Leiden University.
- Thompson, Laurence C. 1965. *A Vietnamese grammar*. Seattle: University of Washington Press.
- Voorhoeve, C. L. 1989. "The masked bird: Linguistic relations in the Bird's Head area." *Peoples on the Move: Current Themes of Anthropological Research in New Guinea* ed. by P. Haenen and J. Pouwer, 78-101. Nijmegen: University of Nijmegen.
- Voorhoeve, C. L. 1994. "Contact-induced language change in the non-Austronesian languages in the north Moluccas, Indonesia." In *Language Contact and Change in the Austronesian World* ed. by T. Dutton and D.T. Tryon, 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.
- Wurm, Stephen A., C.L. Voorhoeve and Kenneth A. McElhanon. 1975. "The Trans-New Guinea Phylum in general." *New Guinea Area Languages and Language Study*, vol. I, *Papuan Languages and the New Guinea Linguistic Scene*, ed. by S.A. Wurm, 299-322. (Pacific Linguistics C-38). Canberra: Australian National University.
- 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.