

# Origins of Numerals and their Variable Forms in the Languages of Pulau Simeulue and Pulau Banyak, Aceh, Indonesia

Asian Linguistic Anthropology  
2022, 4(1): 1-28  
(c) JALA 2022  
Article reuse guidelines  
Jala.pub/permissions  
DOI: 10.47298/jala.v4-i1-a1  
Jala.pub

Robert Amery  
The University of Adelaide, Australia

Zulfadli A. Aziz  
Universitas Syiah Kuala, Banda Aceh, Indonesia



## Abstract

Enumeration in the languages of Pulau Simeulue and Pulau Banyak is a complex topic. A language survey conducted by the authors in 2016 and subsequent investigations revealed a plethora of different forms of the basic numerals from one to ten in each of the four indigenous languages, as well as in Nias spoken by newcomers from the south, depending on what is being counted. Variation is evident not only between the five languages under study, but also within each of the five languages. The variable forms arise from several independent sources. Some of the variation is deep-seated and is also found in the Formosan languages of Taiwan. Some is due to fusion of the numeral and the classifier for 'fruit' to form a new numeral form. However, much of the variation is due to language contact with Jamee and Indonesian. Some minor variants may best be regarded as 'doublets' (Blust, 2013: 278; 338-340). Other variants are explained by the presence or absence of a nasal ligature, but the source of some variants remains obscure. The result is a suite of intricate and complex numeral systems which are now under threat due to language shift. In this fluid context there appears to be considerable variation across speakers and even perhaps within the same speaker, though this remains to be thoroughly investigated.

This paper presents the variation we documented through primary investigation over the period 2016 to 2019 and as documented by Ariani (2016). We attempt to trace the

origins of the individual forms of the numerals in use in the languages of Pulau Simeulue and Pulau Banyak, also known as the Barrier Islands.

**Keywords:** Numerals, Austronesian languages, Simeulue, Pulau Banyak, Devayan

## Introduction

Indonesia is a region that has seen intense language and culture contact over millennia, through waves of traders and religions, punctuated at times by accelerated movements of people in response to natural disasters and sometimes conflict. The province of Aceh, as the intersection of the Indian, Australian and Sunda tectonic plates in Western Indonesia, is no exception. The movement of these plates gave rise to the massive 2004 tsunami, with its epicenter just off the island of Pulau Simeulue, between this island and mainland Aceh. Studies (Daly 2015; Nurjanah et al 2017) show that at least 20 tsunamis of that magnitude have occurred in the region over the past seven thousand years.

Pulau Simeulue is the northernmost member of a chain of islands, known as the Barrier Islands, which lie off the west coast of Sumatra. Pulau Simeulue and Pulau Banyak, the most northerly of the Barrier Islands, are located in Aceh Province. Several languages, namely Devayan, Sigulai, and Leukon, are endemic to Pulau Simeulue, whilst Haloban is only spoken in the village of Haloban/Asontola on the islands of Pulau Tuangku within Pulau Banyak (see Aziz and Amery 2016). These languages differ markedly to the languages indigenous to the mainland of Aceh, such as Acehese, Gayo, and Alas. Acehese mainlanders cannot understand the languages of Pulau Simeulue and Pulau Banyak without making an effort to deliberately learn these languages.



Figure 1. Map of the region and its languages (Zulfadli 2014)

In addition to these languages, Jamee ('guest' in Acehnese), a language descending from Minang in West Sumatra, first appeared in Aceh with the arrival of Minang refugees from the Padri War in West Sumatra (1805 to 1836) who sought refuge in South and West Aceh, Pulau Banyak, and Pulau Simeulue. The society at large in Aceh welcomed these strongly Islamic Minang immigrants (Tim Balai Bahasa Aceh 2012: 32). Jamee is the main language of commerce in the administrative centres of Sinabang in south Pulau Simeulue and in Teluk Nibung in Pulau Banyak, overtaking Indigenous languages of the northernmost Barrier Islands, where it has gained a major foothold (see Sulyanti 2013). In Pulau Banyak, the Jamee language is widespread throughout, except in Ujung Sialit, where the Nias language is dominant. Furthermore, in Pulau Banyak, immigrants from Nias, a large island to the south in the province of North Sumatra, have settled in two small villages, Ujung Sialit and Suka Makmur. Whilst Suka Makmur is a Muslim village, Ujung Sialit has been settled by Christian immigrants from Pulau Nias. From these commercial and administrative centres, Jamee has spread throughout the region as the first language of most of the population. However, the standard Indonesian language, as the national and official language, serves as the language of administration, of education, and in these regions, the main language of religion. Indonesian and Jamee together thus profoundly impact the Indigenous languages of Pulau Simeulue and Pulau Banyak.

People in Pulau Simeulue and Pulau Banyak exclusively use the four Indigenous languages of this region (Devayan, Sigulai, Leukon, and Haloban), that is, these languages have no presence on the Acehnese mainland. On these islands, the majority of the population only speak these languages plus Bahasa Jamee and some Bahasa Indonesia, and not Acehnese (Zulfadli 2014). Aziz and Amery (2016) observe that the languages spoken in Simeulue and Banyak Islands are closely related, and appear to be both rooted in the Nias language. Especially interesting subsystems of the languages of Pulau Simeulue and Pulau Banyak are the counting system(s) and the use of classifiers. Indeed, there is more than one means of enumeration, a plurality contingent on the nature of the entity being counted.

This present study responds to the question of how the similarities, and differences in the enumeration and classification systems of the languages spoken in Pulau Simeulue and Pulau Banyak have arisen. Based on the results of our language survey and subsequent enquiry, we have found that the systems of enumeration in each of the four languages and in Bahasa Nias are closely related. Whilst there are strong resemblances across the languages, each language exhibits its own particular constellation of variations. As such, we structure the paper as follows: In the second section, we begin with a review of the literature on systems of enumeration and numeral classifiers, with particular emphasis on languages of the region. In the third section, we present our methodological framework and the sources of our data. The fourth section presents and discusses the findings and the sources of variation of language in Pulau Simeulue and Pulau Banyak, whilst Section 5 draws conclusions to the paper, calling for more detailed and thorough investigations of this interesting and complex topic.

## Literature Review

Many systems of enumeration appear in the world's languages. Numeral systems typically make use of a base from which to construct their numeral expressions (Schapper and Klamer 2014). According to Comrie (2013), the base of a numeral system is the value  $n$  from the pattern  $xn + y$  where some numeral  $x$  is multiplied by the base plus some other numeral. Schapper and Klamer (2014) argue that many languages have multiple bases, such as Dutch numerals which have five different bases: *tien* '10,' *honderd* '100,' *duizend* '1000,' *miljoen* '100,000.' Papuan languages, for example, have body-part tally systems and restricted numeral systems which have no cyclically recurring base (Schapper and Klamer 2014). Klamer and Kratochvil (2014: 3), cited by Blust (2009: 268-282), Comrie (2013), and Schapper and Hammarström (2013), point out that in Austronesian languages, decimal systems mostly predominate, even though other possibilities are also attested: Innovative forms with ternary and quinary bases; complex numeral formations which involve additive, subtractive, and multiplicative procedures; and systems that mix various types of bases.

Despite considerable variation in the system of enumeration as we summarize above, numerals themselves are usually invariant. In English, for instance, 1 item is 'one,' 2 items are 'two,' and 3 items are 'three,' irrespective of the item being counted, though alternative words such as 'dozen,' 'score,' or 'century,' may replace 'twelve,' 'twenty,' or 'one hundred,' respectively, predicated on context. This phenomenon also occurs in the vast majority of the world's languages, including Bahasa Indonesia. The languages of Pulau Simeulue and Pulau Banyak are quite unremarkable in relation to the system itself; They all employ a base-10 counting system as do Indonesian and English. However, the form of the numeral does vary considerably, depending on the item being counted.

Southeast Asian Languages have an extensive range of number systems for quantification and measurement. Aikhenvald (2000: 1) proposes that "almost all languages have some grammatical means for the linguistic categorization of nouns and nominals," where

different types of classifier can be distinguished by their grammatical status, degree of grammaticalization, conditions for use, meaning, kinds of origin, mode of acquisition, and tendencies towards loss.

As words through which to signify and calculate the number of tangible entities (e.g., people, animals, or goods) and concepts (Alwi et al. 2003: 275), some categories of numerals contain certain characteristics (Kridalaksana 2007: 79), such as succeeding nouns in syntactic construction, and their inability to join with *tidak* 'not' and *sangat* 'very,' such as *seekor anjing* 'a dog' but not *\*tidak seekor anjing* or *\*sangat seekor anjing*.

Some languages use nouns as classifiers, as in Yidiny, an Australian language (Dixon 1982: 192) in the phrase *bama waguja*, where the classifier *bama* 'CLPERSON' co-occurs

with *waguja* ‘man’ to mean ‘a man.’ Other languages, Aikhenvald (2000: 2) notes, have certain words or morphemes which usually occur next to a numeral, or a quantifier, and these words or morphemes “may categorize the referent of a noun in terms of its animacy, shape, and other inherent properties.” Such words or morphemes are referred to as numeral classifiers. Languages with noun classifiers are characterized by the co-occurrence of another noun in a noun phrase. However, Aikhenvald (2000: 98) claims that the numeral classifiers may be the most commonly known type of classifier system, and that they usually appear “contiguous to numerals in numeral noun phrases and expressions of quantity.”

Generally, we divide numerals into two categories; cardinals, which simply present numbers, and ordinals, which indicate a sequence (Quirk et al. 1985: 251-252). In Bahasa Indonesia, for example, ordinal numerals are usually formed through affixation of numerical base morphemes with the prefix *ke-* (Retnaningsih 2015: 2), such as *kedua* ‘second’ and *ketiga* ‘third.’ Retnaningsih suggests that more complex forms are also derived by compounding, conversion, and reduplication. Alwi et al. (2003) have classified numerals in Bahasa Indonesia into eight types; definite numerals, indefinite numerals, distributive numerals, collective numerals, fractional numerals, numerals level, measurement numerals, and classifier numerals. In the Devayan language, spoken on Simeulue island, Ariani (2016) has also found the same eight types, which are to some extent similar to Bahasa Indonesia in terms of their usage. Of particular interest in Ariani’s paper are the definite numerals and their co-occurrence with classifiers of different kinds.

Definite numerals comprise the base or core numerals 0 (zero) to 9 (nine), which serve as the units, and which are then followed by tens, hundreds, thousands, millions, and billions, consecutively. The base numerals combine with the aforementioned forms (tens, hundreds, thousands, etc) to enable a continuous sequence of numerals from zero to the billions and beyond (Alwi et al. 2003: 276).

The numeral level, for example, ‘second,’ ‘third,’ etc, is a transformation from the main numeral (Khallifitriansyah 2002: 60). In English these numerals are derived from the base numeral by means of affixation (-th, -st, -rd) and suppletion, and refer to multiples which indicate the frequencies or the relative positioning of events. In Indonesian, these terms can be formed by adding the prefix *ke-* numbers. The functions of this numeral include to declare the noun sequence, the number of events, and the time sequence. For example, *kelima* ‘fifth,’ is formed from the base form *lima* ‘five’ with the prefix *ke-* attached. Chaer (2006: 115) proposes three rules for numeral levels in Indonesian: Firstly, when the numeral intends to represent ‘place sequence,’ it appears behind the noun, as in Example 1 below:

<i>Adikku menjadi juara keempat lomba baca puisi</i>
--

'My sister became the fourth champion in the poetry reading contest.'
---

Example 1 (Bahasa Indonesia)
------------------------------

Secondly, in order to represent 'sequence or quality and quantity order' within the numeral level, an absorption number is used, such as *primer* 'primary,' *sekunder* 'secondary,' *tertier* 'tertiary,' etc. See for instance, Example (2) below:

<i>Beras merupakan kebutuhan primer masyarakat Indonesia</i>
--

'Rice is a primary necessity in Indonesia,'
---

Example 2 (Bahasa Indonesia)
------------------------------

In Example 2 above, the speakers suggests that rice is sequentially the first need in the country.

Thirdly, when the numerical level aims to represent the set number, it appears before the noun, as in Example 3 below:

<i>Kedua anak itu datang terlambat.</i>
---

'Those two kids come late.'
-----------------------------

Example 3 (Bahasa Indonesia)
------------------------------

This latter function of the Indonesian prefix *ke-* is not served by the suffix *-th* or any overt marking in English.

Measurement numerals signify size or measurement. According to Alwi et al. (2003: 280), measurement numerals signify the size of weight, length, or amount. For example in Indonesian, the measurement words *lusin* 'dozen,' *liter* 'liter,' or *gram* 'gram' can be preceded by numerals as in Example 4 below:

<i>Kalau ketoko, belilah dua lusin piring.</i>
--

'if you go to the store, buy two dozen plates.'
---

Example 4 (Bahasa Indonesia)
------------------------------

Numeral classifiers represent nouns with auxiliary word number, which Chaer (2008: 94) argues are also called match auxiliary number. Similarly, numeral classifiers are words that act as noun identifiers and are placed between the numerals and the noun.

They are used to express the unit number of single objects such as *orang* 'person,' *ekor* 'tail,' *buah* 'fruit,' and other objects. Aikhenvald (2000: 98-99) informs that numeral classifiers are usually in the form of independent lexemes, but can sometimes be affixed to numerals, as in Example 5 below.

<i>ek-ti bai</i>
'one beautiful small book'
Example 5 (Bengali)

Classifiers such as *gelas* 'glass' or *sendok* 'spoon' as in *dua gelas air* 'two glasses of water' or *empat sendok gula* 'four spoons of sugar' appear much more often in Indonesian than in English (Newnham and Wackett 2006). According to Sneddon et al. (2010), the classifiers in Indonesian are always preceded by a number, and when the number plus classifiers appear, they can follow the noun. When classifiers appear with nouns, the nouns are always indefinite. Sneddon et al. (2010) also suggest that in the Indonesian language, there are only three classifiers which are mostly frequently used: *orang* 'person,' *ekor* 'tail,' and *buah* 'fruit.' In Acehnese, *ureueng* corresponds to *orang* 'person' in Indonesian but it is not used as a classifier. Rather, *droe* 'self' acts as a special classifier for humans. All the other classifiers are confined to non-humans (see Durie 1985: 139), such as *boh* 'fruit' which can be used for both fruits and/or animals.

*Orang* 'person' and *ekor* 'tail' appear as classifiers in Indonesian and refer to human beings and animals respectively. These two classifiers directly follow the number yet precede the noun, for example,

<i>tiga ekor harimau</i>
three tail/unit tiger
'three tigers'
Example 7 (Bahasa Indonesia)

<i>lima orang laki-laki</i>
five person male(pl)
'five men.'
Example 7 (Bahasa Indonesia)

We should note, however, that the prefix *se-* attached to the classifier is mostly used instead of *satu* ‘one,’ for example *seekor* ‘one animal’ or *seorang* ‘a person, one person’ (see Sneddon et al. 2010: 138; Newnham and Wackett 2006: 92-93). The other classifier that appears very often in Indonesian is *buah* ‘fruit.’ *Buah* classifies most inanimate objects or things, especially large and/or round objects.

In addition to the three most frequently used Indonesian classifiers, many others exist but are becoming obsolete (Sneddon et al. 2010). Some classifiers overlap in function. These include: *helai* ‘piece,’ *lembar* ‘sheet,’ and *carik* ‘piece,’ which can all co-occur with *kertas* ‘paper;’ *biji* ‘seed,’ and *butir* ‘grain,’ sometimes overlapping in their use with *telor* ‘egg;’ *batang*, which is used with cylindrical objects such as pipes, tree trunks and cigarettes; *bentuk*, which is used for round and curved objects; *bilah*, which is used for sharp things such as knives and needles; and *bidang*, which is used with flat, spread-out items, as in *sebidang tanah* ‘a plot of land.’

According to Marnita (1996), in the Minang variety of the Malay language spoken in West Sumatra, numeral classifiers are obligatory with numerals one to three, but are optional with others. This language contains noun classifiers and a set of numeral classifiers. Consider the examples below (Gil, 2013):

<i>dua ikue anjiang</i>
two CL dog
‘two dogs’
Example 8 (Minang)

<i>tigo batang pituluik</i>
three CL pencil
‘three pencils’
Example 9 (Minang)

<i>buah limau</i>
CL:FRUIT lemon
‘lemon fruit’
Example 10 (Minang)

<i>batang limau</i>
CL:TREE lemon
'lemon tree'
Example 11 (Minang)

The construction of numeral classifiers certainly varies from one language to another. This variation and use are usually determined by the constituent order and constituency relations in classifier constructions that depend on the general syntactic rules of the language (Greenberg 1963; 1972; Dryer 1992). Greenberg (1972), as cited in Aikhenvald (2000: 104-105), revealed four possible constituent orders in numeral classifier constructions: [NUM-CL]-N (as in Chinese, Vietnamese, Hmong, Miao of Wei Ning, Uzbek, Hungarian), N-[NUM-CL] (in Thai, Khmer, Mal) [CL-NUM]-N (as in Ibibio), and N-[CL-NUM] (as in Bodo). In Bahasa Indonesia, Minang or Jamee language, and Acehnese, the constituent order in the numeral classifier construction is [NUM-CL]-N.

### Methodological Framework

We developed this study from data first elicited by Amery in 2016 from two proficient speakers of Haloban living in the village of Asantola, Pulau Tuangku, Pulau Banyak, in the province of Aceh. The documentation entailed voice recordings of an elicitation session undertaken as part of a general language survey. These initial findings were followed in 2017-2019 by a more detailed investigation of male and female students from Pulau Simeulue and Pulau Banyak studying at three universities in Banda Aceh and Aceh Besar (Greater Aceh) by Aziz. As such, Aziz selected two male and two female speakers from each language.

We structured the interviews carried out during fieldwork through the use of a written interview guide or protocol, and extracted our data from the interview transcripts. We then analyzed the data, through a comparative description, in three stages, as suggested by Miles et al. (2014); data condensation, data display, and conclusion drawing. In the condensation stage, we grouped the data into enumeration and classifier categories. In the data display stage, we tabulated this collected data to thus assist us to significantly make sense of our findings. Finally in the conclusion drawing stage, we constructed developmental hypotheses of the data, while attempting to discuss this data in narrative form.

During our survey of the languages of Pulau Simeulue and Pulau Banyak in 2016, we became aware of the use of a variety of numerals for telling the time (e.g. *jam duo* 'two o'clock') and quantifying hours (e.g. *dua jam* 'two hours') vs quantifying the number of clocks (e.g. *domba jam* 'two clocks') in Bahasa Devayan. We noted a similar phenomenon

in the literature for Tagalog (Woods 2011). Amery later noticed in his investigation of Haloban in Pulau Banyak, that the system of enumeration mysteriously exhibits a complex interplay between numerals and classifiers. These initial findings led to further investigations by Aziz with speakers of each of the languages of Pulau Simeulue and Pulau Banyak living in Banda Aceh.

As Bahasa Indonesia is well-documented (Sneddon 1996; Chaer 2012; Kridalaksana 2007), it makes sense to use Bahasa Indonesia as the point of reference for a comparison of enumeration in the local languages with enumeration in the national language. We also deemed it significant to compare enumeration in the languages under study with Bahasa Jamee and its predecessor Bahasa Minang, owing to the fact that, along with Bahasa Indonesia, these are the main languages in contact with the languages of Pulau Simeulue and Pulau Banyak. Furthermore, we have chosen to also compare these languages with Acehnese, the dominant language of the province of Aceh, and other languages of the region.

## Findings

The languages of Pulau Simeulue and Pulau Banyak have complex systems of enumeration, drawing on the resources of indigenous counting systems and the more recently introduced Jamee numerals. Jamee numerals are identical or nearly identical to Minangkabau numerals which incidentally overlap significantly with the Indonesian enumeration system. Minangkabau is spoken in the province of Sumatera Barat (West Sumatra), the ancestral homeland of present-day Jamee speakers. Haloban appears to be most heavily influenced by Jamee, where speakers employ Jamee numerals during the counting process, with the exception of *satu* 'one' which is borrowed from Bahasa Indonesia. Devayan and Leukon use indigenous numbers when counting from one to seven, but revert to Jamee numerals for numbers from eight onwards. Sigulai speakers use indigenous numbers up until nine or ten, and Devayan speakers use *fa* 'ten,' which differs to the Jamee *sapuluah* 'ten.' All four languages use Jamee numbers for eleven onwards when counting, for example *tuju(h) bale* 'seventeen,' even though Sigulai used *fitu* 'seven' and Devayan and Leukon used *itufu* 'seven,' which differ to the Jamee *tujuh* (cf Bahasa Indonesia *tujuh* 'seven'). Table 1 below presents the counting systems for the languages of Pulau Simeulue and Pulau Banyak along with Jamee and Indonesian.

Comparison of the five Pulau Simeulue and Pulau Banyak languages indicates that the most pronounced variation appears in the form of the numbers for one and two. Sigulai is the most innovative with *amba* 'one' and *dumba* 'two' (which differ from the Haloban *satu* 'one' and *dua* 'two, Devayan/Leukon *sao* 'one' and *dufo* 'two,' and Nias *sara* 'one' and *dua* 'two'). The divergent Sigulai numerals *amba* 'one' and *dumba* 'two' appear to have formed from a fusion of the numeral followed by the classifier for fruit where a transitional homorganic nasal 'm' has been inserted between the numeral and classifier. This is evident when we look at the Nias forms for counting fruit; *sam bua* 'one

fruit' and *dom bua* 'two fruits.' Sigulai shows a much closer affinity with Nias than Devayan, Leukon, or Haloban. The Sigulai numerals *olu* 'eight' and *siwa* 'nine' correspond closely to Nias *walu* 'eight' and *siwa* 'nine,' which differ to the Jamee-derived *salapan* 'eight' and *sambelan* 'nine' in all the other languages. The Sigulai numerals 'one' to 'ten' in Table 1 all use the same roots as the Nias roots with some minor phonological changes, such as loss of the initial s- in *amba* 'one,' loss of the initial w- in the *olu* 'eight,' loss of the final -h in the *fuluh* 'ten,' and vowel changes ö -> e in the *efa* 'four,' the *ene* 'six' and a -> o in the *olu* 'eight.'

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	Sao	Amba	Sao	Satu	Sara	ciek	Satu
2	Dufo	Dumba	Dufo	Dua	Dua	duo	Dua
3	Tellufo	Telu	Tlufo	Tigo	Tölu	tigo	Tiga
4	Atao	Efa	Atafo	Ampek	Öfa	ampek	Empat
5	Limafo	Lima	Limafo	Limo	Lima	limo	Lima
6	Nemmafo	Ene	Nemmafo	Anam	Önö	anam	Enam
7	Itufo	Fitu	Itufo	Tujuh	Fitu	tujuh	Tujuh
8	Salapan	Olu	Salapan	Lapan	Walu	salapan	Delapan
9	Sambelan	Siwa	Sambélan	Sambilan	Siwa	sambilan	Sembilan
10	Faan	Fuluh	Sapulu	Sapuluh	Fulu	sapuluah	Sepuluh

Table 1: Counting systems in Pulau Simeulue and Pulau Banyak languages

The Devayan and Leukon numerals 'one' to 'seven' may have developed in a similar way. The suffixed element in these numerals appears to have three forms; -o, -fo and -afo, all of which have probably also arisen from a fusing of the classifier *bo* ~ *fo*1 'fruit.' The allomorph -afo occurs on the consonant final root *nem* 'six.' However, the object or element that conditions the -o allomorph on 'one' in Devayan and Leukon and on 'eight' in Devayan remains ambiguous.

The ordinal numbers in the five languages spoken in the islands, with the exception of Nias, are surprisingly all derived from the Jamee language. It seems that Devayan, Sigulai, Leukon, and Haloban have all adopted the Jamee ordinals, with only slight differences. In most cases, the roots of the ordinals exactly match those in Jamee, with the exception of seven and ten, where *u* or *uh* replace the distinctive *uah* ending. However, the prefixes *per-* and *ke-* are derived from Indonesian, and replace the Jamee forms *par-* and *ka-*. Nias, however, has maintained its conservative forms.

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee
First	Pertamo	Pertamo	Pertamo	Pertamo	sara	Partamo
Second	Keduo	Keduo	Keduo	Keruo	dua	Kaduo
Third	Ketigo	Ketigo	Ketigo	Ketigo	tölu	Katigo
Fourth	Keampek	Keampek	Keampek	Keampek	öfa	Kaampek
Fifth	Kelimo	Kelimo	Kelimo	Kelimo	lima	Kalimo
Sixth	Keanam	Keanam	Keanam	Keanam	önö	Kaanam
Seventh	Ketuju	Ketujuh	Ketuju	Ketujuh	fitu	Katujuah
Eighth	Kelapan	Kelapan	Kelapan	Kelapan	walu	Kalapan
Ninth	Kesambilan	Kesambilan	Kesambilan	Kesambilan	siwa	Kasambilan
Tenth	Kesapulu	Kesapuluh	Kesapulu	Kesapuluh	fulu	Kasapuluah

Table 2: Ordinal Numbers in Pulau Simeulue and Pulau Banyak languages

Table 2 maps the affinity among the languages Devayan, Sigulai, Leukon, Haloban, and Jamee, whilst Nias uses its own forms, similar to the cardinal numbers in Table 1 above.

The pronunciation of the ordinal numbers ‘seventh’ and ‘tenth’ also differs in Devayan and Leukon, which omit the word-final /h/ when spoken. Only Nias has maintained the conservative forms of its ordinal numbers, which are identical to cardinal numerals. In Haloban, the ordinals derive from the cardinals, but the roots across Devayan, Sigulai and Leukon (the languages of Pulau Simeulue), differ in most cases.

In counting various types of objects, the form of the numeral changes as a result of three factors; the fusing of the numeral with the classifier, phonological changes, or choosing between the indigenous or introduced (Jamee or Indonesian) numeral. We consider the example of Sigulai; *amba* ‘one,’ *mea* ‘one’ (person), *alangaikha* ‘one’ (animal), *ageu* ‘one’ (long square or round object), *ala* ‘one’ (long flat object eg *ala biuk* ‘one knife’), *amba buah* ‘one fruit’ *amba* ‘one’ (small container), *ambebe* ‘one bound object’ (from *amba* + *bebe*), *amba bine* ‘one package,’ *sakilo* ‘one kilo,’ *amba* ‘one’ (seed-like object), *sameter* ‘one metre,’ *amba* ‘one’ (bottle, plate etc), *avilang* ‘one section’ (of land, road etc), *akhete* or *alabulu* ‘one fine, thin object’ (*khete* ‘stalk,’ *bulu* ‘leaf’) and so on. Here, *amba*, *ala*, *a-*, *am-*, *sa-* and arguably  $\emptyset$  vary in the element meaning ‘one.’ In *alangaika* ‘one’ (animal), an epenthetic *nga* syllable or ligature emerges between the numeral and classifier. This same ligature appears very often in numerals in Devayan, Sigulai, and Leukon, and more often in the larger numerals.

No.	People	Animals	Fruits	Thin flat objects	Firewood bundles	Long cylind. objects
1	Mesa	sara	Sao	Salemba	Saferet	Sangaon
2	Daro	Dua	Dufo	Duo lembar	Doferet	Dongaon
3	Datello	Tlosiha	Tellufo	Tigo lembar	Tlo feret	Tlongaon
4	Dahat	Akasiha	Atao	Ampek lembar	Aka feret	Akangaon
5	Dallima	Limasiha	Limafo	Limo lembar	Limo feret	Limangaon
6	Daennem	Nemengasiha	Nemmafo	Annam lembar	Nemenga feret	Nemengangaon
7	Daitu	Itungasiha	Itufo	Tuju lembar	Itunga feret	Itungangaon
8	Salapan	Salapanngasiha	Salapan	Salapan lembar	Salapannga feret	Salapanngangaon
9	Sambelan	Sambelanngasiha	Sambelan	Sambelan lembar	Sambelannga fere	Sambelanngangaon
10	Sapulu	Sapulungasiha	Faan	Sapulu lembar	Sapulunga feret	sapulungangaon

Table 3: Enumeration of six categories in Devayan

As Table 3 presents, in Devayan, the various forms of the numeral ‘one’ are best explained through the presence of a root *sara*,<sup>2</sup> which is reduced to *sa-* when combined with a classifier. The full form *sara* appears without a classifier when counting ‘animals.’ When enumerating one person, a prefix *me-* combines with the reduced form *sa*. When counting the other four entities, fruits, thin flat objects, bundles of firewood, and long cylindrical objects, *sa-* precedes a numeral classifier (*-o* ‘fruit,’ *lembar* ‘sheet’ for ‘thin flat objects,’ *feret* for ‘bundles of firewood’ and *ngaon* for ‘long cylindrical objects’) in the same way that *sa-* prefixes the numeral classifier in Jamee and *se-* in Indonesian. The Indonesian numeral *dua* ‘two’ appears without a classifier when counting animals, in contrast to Indigenous numerals for counting from three to seven animals.

The numeral for counting ‘people’ in Devayan has its own characteristic prefix. Apart from *mesa* ‘one’ (person), the numbers two to seven contain a prefix *da-*, whilst the counting of eight, nine, or ten people requires Jamee-derived numerals without the prefix *da*. Jamee-derived numerals also appear consistently, together with the Indonesian classifier *lembar*, when counting ‘thin flat objects.’ Devayan consistently requires Jamee-derived numerals to appear in cases of counting more than seven entities, except for *faan* ‘ten’ (fruits). A ligature *-nga* appears between the numeral and classifier, in the case of numerals six and above, but only when counting ‘animals,’ ‘bundles of firewood,’ and ‘long cylindrical objects.’ There is some further variation in the form of the Indigenous roots (*ro ~ du ~ do* ‘two’; *tello ~ tlo ~ tellu* ‘three’; *hat ~ aka ~ ata* ‘four’; *ennem ~ neme ~ nemma* ‘three’).

No.	People	Animals	Fruits	Thin flat objects	Firewood bundles	Long cylind. objects
1	Mea	Alangaikha	Amba	Salemba	saferet	Ageu
2	Dalua	Duangaikha	Domba	Duo lembar	Doferet	Dogeu
3	Datelu	Telungaikha	Telu	Tigo lembar	Tlo feret	Telungageu
4	Daeva	Evangaikha	Eva	Ampek lembar	Aat feret	evangageu
5	Dalima	limangaikha	Lima	Limo lembar	Limo feret	Limangageu
6	Daena	Enengaikha	Ene	Annam lembar	Nem feret	enengageu
7	Dafitu	Fitungaikha	Fitu	Tujuh lembar	Itu feret	fitungageu
8	Daolu	Olungaikha	Olu	Lapan lembar	Salapan feret	olungageu
9	Dasiwa	Singangaikha	Siwa	Sambelan lembar	Sambelan feret	siwangageu
10	Dafulu	Fulungaikha	Fuluh	Sapuluh lembar	Sapulu feret	fulungageu

Table 4: Enumeration of six categories in Sigulai

Table 4 shows Sigulai as the most conservative of the Pulau Simeulue and Pulau Banyak languages, with Indigenous roots entirely when enumerating ‘people,’ ‘fruit,’ and ‘long cylindrical objects,’ mostly when counting ‘animals,’ except for number two and for two to seven ‘bundles of firewood.’ A variation appears in some of these indigenous roots, though this is less than in Devayan and Leukon. For instance, the indigenous root for ‘six’ is variably *ena* ~ *ene* ~ *nem*, which employ Jamee forms for one and eight to ten ‘bundles of firewood,’ and for one to ten ‘thin flat objects.’ The numerals for counting fruit are identical to those for counting in the absence of any specific entity (see Table 1). Similarly, no classifier appears, except for the bleached classifier, in the numbers *amba* ‘one’ and *domba* ‘two’ discussed earlier following Table 1. Like Leukon (Table 5), Sigulai requires the ligature *-nga* when counting ‘animals’ and more than two ‘long cylindrical objects.’ In common with the other languages, Sigulai has a prefix *da-* when counting ‘people.’

No.	People	Animals	Fruits	Thin flat objects	Firewood bundles	Long cylind. objects
1	Mèsa	Sara Ngasikha	Sao	Sao	saferet	Saayu
2	Daro	Duugasikha	Dufo	Dufo	Doferet	Doayu
3	Datəlô	Tlungasikha	Tlufo	Tlufo	Tlo feret	Teloayu
4	Dahat	Aat/akangasikha	Atafo	Atafo	Aat feret	Aat ngaayu
5	Dalima	Limangasikha	Limafo	Limafo	Limo feret	Limangayu
6	Dannəm	ənəmngasikha	Nəmmafo	Nəmmafo	Nem feret	ənəm ngayu
7	Daitu	Itungasikha	Itufo	Itufo	Itu feret	Itungayu
8	Salapan	Salapanngasikha	Salapan	Salapan	Salapan feret	salapan ngayu
9	Sambélan	Sambélanngasikha	Sambélan	Sambélan	Sambelan feret	Sambélan ngayu
10	Sapulu	Sapulungasikha	Sapulu	Sapulu	Sapulu feret	Sapulu ngayu

Table 5: Enumeration of six categories in Leukon

Leukon enumeration (Table 5) is very similar to Devayan (Table 3), although minor differences appear in numeral roots. This phenomenon is most pronounced in the numeral six in Devayan, that is, *ennem* ~ *neme* ~ *nemma* ~ *annam*, versus Leukon *nnəm* ~ *ənəm* ~ *nəmma* ~ *nem*). The only major difference in the roots themselves is the use of *faan* ‘ten’ (fruits) in Devayan, whereas Leukon uses *sapulu* ‘ten’ in all counting. Still, Devayan also employs *sapulu* ‘ten’ when counting other entities. The Leukon classifier for long cylindrical objects is *ayu*, whereas in Devayan it is *ngaon*, whilst the Leukon classifier for animals *sikha* only slightly differs to its counterpart *siha* in Devayan. Leukon uses the ligature *nga* less in counting long cylindrical objects, and omits this ligature entirely in counting bundles of firewood, but employs it consistently with all numerals when counting animals. Devayan, however, only employs the ligature *-nga* when counting more than two ‘animals’ or more than five ‘bundles of firewood’ and ‘long cylindrical objects.’ In Leukon, the enumeration of ‘thin flat objects’ is exactly the same as the enumeration of ‘fruit,’ whereas Devayan requires Jamee-derived numerals with the classifier *lembar*, borrowed from Bahasa Indonesia, to count ‘thin flat objects.’

No.	People	Animals	Fruits	Thin flat objects	Firewood bundles	Long cylind. objects
1	mesa	sara	Sao	Salembar	Sewirit	Sangawun
2	raro	rua	Ruo	Duo lembar	Dua wirit	Rongawun
3	ratelu	telu	Teluo	Tigo lembar	Tiga wirit	Teloawun
4	rahat	aka	Atawo	Ampek lembar	Ampat wirit	Akanawun
5	ralima	lima	Limaho	Limo lembar	Lima wirit	Limangawun
6	ranem	enam	Namaho	Annam lembar	Anam wirit	Anamawun
7	raitu	itu	Ituwo	Tujuh lembar	Tujuh wirit	Tujuhawun
8	salapan	lapan	Salapan	Lapan lembar	Salapan wirit	Salapanawun
9	sambilan	sambilan	sambilan	Sambelan lembar	Sambilan wirit	Sambilanawun
10	sapuluh	sepuluh	Sapuluh	Sapuluh lembar	Sapuluh wirit	Sapuluhawun

Table 6: Enumeration of six categories in Haloban

Table 6 above presents Haloban as sharing many numeral classifiers with Devayan and Leukon. The difference in these languages is largely phonological. In Devayan, Sigulai, and Leukon, the prefix *da-* attaches to the numerals two to seven when enumerating people. In Haloban, the prefix *ra-* is used with the same distribution resulting from *d > r* lenition. The suffixed ‘fruit’ classifier *-fo* in Devayan and Leukon has changed to *-o* ~ *-ho* ~ *-wo* in Haloban. In addition, in the ‘bundles of firewood’ numeral, *sa-* changes to *se-* (*wirit*), as in the expression *sewirit ayu* ‘one bundle of firewood,’ but precisely the opposite sound change appears in the numerals *ampat* ‘four’ and *anam* ‘six.’ When Indonesian numerals, eg *sembilan* ‘nine’ or *sepuluh* ‘ten’ are used, they undergo this same *e > a* phonetic change in Haloban.

No.	People	Animals	Fruits	Thin flat objects	Firewood bundles	Long cylind. objects
1	Samesa niha	Sara na'eu	Sam bua	Sara	Saraboloe	Sagewu
2	Darua niha	Dua na'eu	Dom bua	Dua	Duaboloe	Duangaewu
3	Datelu niha	Telu na'eu	Telö ngawua	Telu	Teluboloe	Telungaewu
4	Daefa niha	Efa na'eu	Efangawua	Efa	Eva boloe	Efangaewu
5	Dalima niha	Lima na'eu	limangawua	Lima	Lima boloe	Limangaewu
6	Daönö niha	Önö na'eu	önö ngawua	Önö	Önö boloe	Önöngaewu
7	Dafitu niha	Fitu na'eu	Fitu ngawua	Fitu	Fituboloe	Fitungaewu
8	Dawalu niha	Walu na'eu	Walungawua	Walu	Waluboloe	walungaewu
9	Dasiwa niha	Siwa na'eu	Siwa ngawua	Siwa	Siwaboloe	Siwangaewu
10	Dafulu niha	Fulu na'eu	Fulu ngawua	Fulu	Fulu boloe	Fulungaewu

Table 7: Enumeration of six categories in Nias

Nias (Table 7) displays an easily comprehensible numerical arrangement, as it is much more systematic and not so different from its basic numerals. Numeral classifiers, either as independent words or suffixed elements, appear consistently for all categories in Table 7, except for the cases with 'thin flat objects.' However, similar to the other languages, small differences in the pronunciation of the roots in Nias appear. *Sara* 'one' (the basic word) sometimes contracts to a prefix *sa-* (corresponding to Indonesian *se-*). In counting 'fruit,' an additional *m* appears in phonological anticipation of the classifier *bua* in *sam bua* 'one' and *dom bua* 'two.' A familiar *d > r* lenition occurs in the root *dua > rua* when used with 'people.' The root for 'three' alternates between *telu* and *telö*, and the root for 'four' alternates between *efa* and *eva*. The ligature *-nga* appears between the numeral for two and above when counting 'long cylindrical objects,' and for three and above when counting 'fruit.' Table 8 below compares the means of enumerating people across the five languages plus Jamee and Indonesian which provide points of reference.

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	Mesa	Mea	Mèsa	mesa	Samesa niha	saurang	seorang
2	Daro	Dalua	Daro	raro	Darua niha	Duo urang	Dua orang
3	Datelo	Datelu	Datəlô	ratelu	Datelu niha	Tigo urang	Tiga orang
4	Dahat	Daeva	Dahat	rahat	Daefa niha	Ampek urang	Empat orang
5	Dalima	Dalima	Dalima	ralima	Dalima niha	Limo urang	Lima orang
6	Daennem	Daena	Dannəm	ranem	Daönö niha	anam urang	Enam orang
7	Daitu	Dafitu	Daitu	raitu	Dafitu niha	tujuh urang	Tujuh orang
8	Salapan	Daolu	Salapan	salapan	Dawalu niha	salapan urang	Delapan orang
9	Sambelan	Dasiwa	Sambélan	sambilan	Dasiwa niha	sambilan urang	Sembilan orang
10	Sapulu	Dafulu	Sapulu	sapuluh	Dafulu niha	sapuluah urang	Sepuluh orang

Table 8: Enumeration of 'people' in Simeulue and Pulau Banyak languages

The languages of Pulau Simeulue and Pulau Banyak, including Nias, display the unique form *mesa*, or *mea* in the case of Sigulai, for ‘one,’ when counting people. Perhaps *me-* was historically a prefix of some kind or a reduced nominal, but it bears no resemblance to *ata* ‘person’ used today. From Table 8 above, we see that a numeral classifier for ‘people’ does not appear in Devayan, Sigulai, Leukon, and Haloban. By contrast, Nias uses a classifier *niha* in addition to the prefix *da-*, whilst Jamee and Indonesian use the classifiers *urang* and *orang* respectively, meaning ‘person.’ Another interesting phenomenon that appears in the data in Table 8 is the fact that numerals borrowed from Jamee appear for eight, nine, and ten, in Devayan, Leukon, and Haloban. It is difficult to determine factors that motivate the use of these Jamee-derived forms. According to our native speaker informants, speakers use such forms natively. Similarly, Table 9 below compares the means of counting animals across these same languages.

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	sara	Alangaikha	Sara Ngasikha	sara	Sara na’eu	saekor	seekor
2	Dua	Duangaikha	Duangasikha	rua	Dua na’eu	Duo ekor	Dua ekor
3	Tlosiha	Telungaikha	Tlungasikha	telu	Telu na’eu	Tigo ekor	Tiga ekor
4	Akasiha	Evangaikha	Aat/akangasikha	aka	Efa na’eu	Ampek ekor	Empat ekor
5	Limasiha	limangaikha	Limangasikha	lima	Lima na’eu	Limo ekor	Lima ekor
6	Nemengasiha	Enengaikha	ənəmngasikha	enam	Önö na’eu	anam ekor	Enam ekor
7	Itungasiha	Fitungaikha	Itungasikha	itu	Fitu na’eu	tujuh ekor	Tujuh ekor
8	Salapanngasiha	Olungaikha	Salapanngasikha	lapan	Walu na’eu	salapan ekor	Delapan ekor
9	Sambelanngasiha	Singangaikha	Sambélanngasikha	sambilan	Siwa na’eu	sambilan ekor	Sembilan ekro
10	Sapulungasiha	Fulungaikha	Sapulungasikha	sepuluh	Fulu na’eu	sapuluah ekor	Sepuluh ekor

Table 9: Enumeration of ‘animals’ in Simeulue and Pulau Banyak languages

In enumerating animals, a classifier (*siha* in Devayan, *ikha* in Sigulai, *sikha* in Leukon, *na’eu* in Nias, and *ekor* ‘tail’ in Jamee and Indonesian) appears in all languages except Haloban. In Devayan, *siha* has no independent meaning for native speakers, and acts simply as a classifier. In Devayan, *iyok* is ‘tail,’ whereas in Nias *i’o* is ‘tail.’ However, *iyok* and *i’o* never appear as in *\*saraiyok* or *\*sara i’o*. This phenomenon also applies to other Pulau Simeulue and Pulau Banyak languages. Therefore, this complexity is one of several unique characteristics in these languages which is unexplainable. In Bahasa Indonesia and Jamee, all numbers co-occur with a classifier, yet this is not the case in Haloban, where animals are enumerated without classifiers. In Haloban, the name of the animal directly follows the numbers. Furthermore, in Devayan, the numeral classifier *siha* for animals does not appear with *sara* ‘one’ and *dua* ‘two.’ For six and above, the ligature – *nga* appears between the numeral and the classifier. In Devayan, Leukon, and Haloban, Jamee-derived numerals appear when counting eight, nine, or ten animals. Table 10 below shows the means of counting fruits across these same seven languages:

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	Sao	Amba	Sao	sao	Sam bua	Sabuah	sebuah
2	Dufo	Domba	Dufo	ruo	Dom bua	Duo buah	Dua buah
3	Tellufo	Telu	Tlufo	teluo	Telö ngawua	Tigo buah	Tiga buah
4	Atafo	Eva	Atafo	atawo	Efangawua	Ampek buah	Empat buah
5	Limafo	Lima	Limafo	limaho	limangawua	Limo buah	Lima buah
6	Nemmafo	Ene	Nemmafo	namaho	önö ngawua	anam buah	Enam buah
7	Itufo	Fitu	Itufo	ituwo	Fitu ngawua	tujuh buah	Tujuh buah
8	Salapan	Olu	Salapan	salapan	Walungawua	salapan buah	Delapan buah
9	Sambelan	Siwa	Sambélan	sambilan	Siwa ngawua	sambilan buah	Sembilan buah
10	Faan	Fuluh	Sapulu	sapuluh	Fulu ngawua	sapuluah buah	Sepuluh buah

Table 10: Enumeration of ‘fruits’ in Simeulue and Pulau Banyak languages

Table 10 shows that the enumeration of ‘fruits’ in Devayan, Sigulai, Leukon, and Haloban remains the same as the system of counting in isolation in each of the languages (see Table 1). However, Nias has its own numeral classifier form for fruits in that its speakers add *bua* or *ngawua* after the numbers. The numeral classifier for ‘fruits’ in Jamee is identical to that in Bahasa Indonesia. Devayan and Leukon have the suffix *-fo* attached to numerals two to seven and *-o* attached to *sa-* ‘one.’ Haloban has a corresponding suffix *-wo ~ -ho ~ -o*, but it is not possible to formulate a phonological rule governing the distribution of the allomorphs. These suffixes seem to have all developed from *bo* ‘fruit’ as a result of lenition. Indeed, lenition is already evident in the Nias forms for ‘three fruits’ and above with the weakening of *b* to *w*.

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	Salembar	Salembar	Sao	Salembar	Sara	salembar	selembar
2	Duo lembar	Duo lembar	Dufo	Duo lembar	Dua	Duo lembar	Dua lembar
3	Tigo lembar	Tigo lembar	Tlufo	Tigo lembar	Telu	Tigo lembar	Tiga lembar
4	Ampek lembar	Ampek lembar	Atafo	Ampek lembar	Efa	Ampek lembar	Empat lembar
5	Limo lembar	Limo lembar	Limafo	Limo lembar	Lima	Limo lembar	Lima lembar
6	Annam lembar	Annam lembar	Nemmafo	Annam lembar	Önö	anam lembar	Enam lembar
7	Tuju lembar	Tujuh lembar	Itufo	Tujuh lembar	Fitu	tujuh lembar	Tujuh lembar
8	Salapan lembar	Lapan lembar	Salapan	Lapan lembar	Walu	salapan lembar	Delapan lembar
9	Sambelan lembar	Sambelan lembar	Sambélan	Sambelan lembar	Siwa	sambilan lembar	Sembilan lembar
10	Sapulu lembar	Sapuluh lembar	Sapulu	Sapuluh lembar	Fulu	sapuluah lembar	Sepuluh lembar

Table 11: Enumeration of ‘thin flat objects’ in Simeulue and Pulau Banyak languages

Differing from other languages in Pulau Simeulue and Pulau Banyak, Leukon and Nias (Table 11 above) maintain the use of the basic counting system by omitting the classifier *lembar* ‘sheet’ after the numbers as in other languages in the islands. Devayan, Sigulai, and Haloban all consistently employ Jamee-derived numerals when counting ‘thin flat objects’, whilst Leukon also employs Jamee-derived numerals for ‘eight,’ ‘nine,’ and ‘ten.’ Note that the ligature *-nga* is entirely absent in Table 11.

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	Saferet	Saferet	Saferet	Sewirit	Saraboloe	Saikat	Seikat
2	Doferet	Doferet	Doferet	Dua wirit	Duaboloe	Duo ikat	Dua ikat
3	Tlo feret	Tlo feret	Tlo feret	Tiga wirit	Teluboloe	Tigo ikat	Tiga ikat
4	Aka feret	Aat feret	Aat feret	Ampat wirit	Eva boloe	Ampek ikat	Empat ikat
5	Limo feret	Limo feret	Limo feret	Lima wirit	Lima boloe	Limo ikat	Lima ikat
6	Nemenga feret	Nem feret	Nem feret	Anam wirit	Önö boloe	anam ikat	Enam ikat
7	Itunga feret	Itu feret	Itu feret	Tujuh wirit	Fituboloe	tujuh ikat	Tujuh ikat
8	Salapannga feret	Salapan feret	Salapan feret	Salapan wirit	Waluboloe	salapan ikat	Delapan ikat
9	Sambelannga feret	Sambelan feret	Sambelan feret	Sambilan wirit	Siwaboloe	sambilan ikat	Sembilan ikat

Table 12: Enumeration of ‘Bundles of firewood’ in Simeulue and Pulau Banyak languages

Table 12 shows the fact that a classifier consistently appears with numerals in all mentioned languages. The classifier *feret* in Devayan, Sigulai, and Leukon, and *wirit* in Haloban, signify ‘tie’ or ‘bundle’ as an independent word. This classifier seems to be unique in Devayan, though the ligature *-nga* appears after the number for six and above, whereas it is fully absent in other languages.

No.	Devayan	Sigulai	Leukon	Haloban	Nias	Jamee	Indonesia
1	Sangaon	Ageu	Saayu	Sangawun	Sagewu	sabatang	sebatang
2	Dongaon	Dogeu	Doayu	Rongawun	Duangaewu	Duo batang	Dua batang
3	Tlongaon	Telungageu	Teloayu	Teloawun	Telungaewu	Tigo batang	Tiga batang
4	Akangaon	evangageu	Aat ngaayu	Akanawun	Efangaewu	Ampek batang	Empat batang
5	Limangaon	Limangageu	Limangayu	Limangawun	Limangaewu	Limo batang	Lima batang
6	Nemengangaon	enengageu	ənəm ngayu	Anamawun	Önöngaewu	anam batang	Enam batang
7	Itungangaon	fitungageu	Itungayu	Tujuhawun	Fitungaewu	tujuh batang	Tujuh batang
8	Salapanngangaon	olungageu	salapan ngayu	Salapanawun	walungaewu	salapan batang	Delapan batang
9	Sambelanngangaon	siwangageu	Sambélan ngayu	Sambilanawun	Siwangaewu	sambilan batang	Sembilan batang
10	sapulungangaon	fulungageu	Sapulu ngayu	Sapuluhawun	Fulungaewu	sapuluah batang	Sepuluh batang

Table 13: Enumeration of ‘Long cylinder-shape objects’ (Simeulue and Pulau Banyak languages)

Table 13 reveals considerable diversity regarding the classifier for ‘long cylindrical objects;’ *ngaon* (Devayan), *geu* (Sigulai), *ayu* (Leukon) *awun* (Haloban), and *gewu/ewu* (Nias). These classifiers differ to Jamee and Indonesian which both use *batang* ‘stick, pole.’ The Devayan equivalent of *batang* is *awak*, and hence not *ngaon*, which has no independent meaning for native speakers of Devayan (pc Siti Nasuha, Aug. 2020). A similar pattern also appears in Sigulai, Leukon, Haloban, and Nias. Note that the ligature *-nga* plays some role in each of the languages of Pulau Simeulue and Pulau Banyak. In Nias, it appears with numerals greater than one. In Sigulai, it appears with numerals greater than two. In Leukon, it appears with numerals greater than three. In Devayan, its use begins with six. Haloban has the strangest distribution of all where it appears only with ‘one,’ ‘two,’ and ‘five.’

## Discussion

The languages of Pulau Simeulue and Pulau Banyak display considerable variation in the shape of their numerals, which also vary according to the object counted. A number of sources influence this variation, including the fact that the original system of the proto-language has been lost in many Austronesian languages, but retained in this particular set of languages. Reflexes of all the indigenous numerals of the languages of Pulau Simeulue and Pulau Banyak appear in the Austronesian languages of Taiwan (Rosenfelder n.d.)<sup>3</sup> and in Proto-Austronesian. They are especially evident in Puyuma, a Paiwanic language from the southeast of Taiwan in the vicinity of Taitung, pointing to their antiquity as opposed to the more recent innovations in many of the Jamee forms that are more closely aligned with Malay and Bahasa Indonesia.

In large part, the variation in these languages has emanated from language contact. The fusion of the subsequent classifier with the numeral to form a new numeral root has also contributed to variation. Concurrently, considerable tinkering with the shape of the numeral has occurred, resulting in minor variations for both indigenous and introduced roots.

### *Independent Numerals ‘One’ and ‘Two’*

The most obvious of all differences across number systems in these languages are the independent numbers signifying ‘one’ and ‘two.’ Devayan/Leukon *sao* ‘one’ vs Sigulai *amba* ‘one’ vs Haloban *satu* ‘one’ (from Indonesian) and Nias *sara* ‘one,’ which is the same form as *sara* in Devayan, Leukon, and Haloban, and also corresponds to Sigulai *ala*, which all appear in the expression for ‘one animal.’ Similarly, Devayan/Leukon *dufo* ‘two’ corresponds to Sigulai *dumba* ‘two’ and Haloban/Nias/Indonesian *dua* ‘two.’ A pertinent concern at present then is the origin of the divergent terms *sao*, *amba*, *dufo*, and *dumba*. The Nias, Jamee, and Indonesian terms for ‘one fruit’ and ‘two fruits’ provide the response to this. The Nias *sam bua* ‘one fruit,’ compared with the Indonesian *sebuah* and Jamee *sabuah*, has introduced a homorganic nasal m in anticipation of the following

b in *bua*. Exactly the same process has occurred in the Nias *dom bua* ‘two fruits,’ compared with the Indonesian *dua buah* and the Jamee *duo buah*. Sigulai *amba* derives from *sambua*: The deletion of the initial *s* is a semi-regular sound change in Sigulai, where the initial *s* is retained in Devayan and Leukon. Similarly, Sigulai *dumba* derives from Nias *dombua*. As such, a process of fusion of *bua* ‘fruit’ (the classifier for fruits) has occurred with the numerals one and two. The classifier has been bleached, devoid of its original meaning, so that *amba* and *dumba* have become independent numerals in their own right, and are no longer restricted to counting in the domain of fruit. The Devayan/Leukon *sao* ‘one’ and *dufo* ‘two’ have evolved in a similar way, where the word for ‘fruit’ in Devayan and Leukon is *bo*.

The use of the word *bo* to classify the numbers in Devayan must be followed by the name of the fruit by adding the phoneme /n/ at the end of the word *bo*, as in the following:

Example 12.1
<i>sao bon mangga</i>
‘one mango’
Example 12.2
<i>dufo bon yamor</i>
‘two papayas’
Example 12.3
<i>tigo bon bonnol</i>
‘three coconuts’
Example Set 12 (Devayan)

According to our Devayan language informant, the addition of the word *bo* without a subsequent name of the fruit will sound strange to its speakers, as in the following:

* <i>sao bo</i> or * <i>sao bon</i> ‘one fruit’ are not possible.
Example 13 (Devayan)

Example 13 further evidences the fact that the final *-o* or *-fo* in Devayan and Leukon numerals, whilst being derived from *bo* ‘fruit,’ are bleached and no longer recognised by native speakers as anything other than a part of the root that cannot be further broken down.

### *Counting People: Prefix da- ~ ra-*

The word *mesa* in Devayan, Leukon, and Haloban (in Amery’s data) or *mea* in Sigulai, signifies ‘one person.’ This word is quite distinct from the independent word for ‘one’

(*sao, amba, sara, satu*). In Aziz's Haloban data, this word *mesa* is combined with *ata* 'person,' forming *mesata* 'one person.' In Nias, *mesa* appears with the additional prefix *sa-* 'one' together with *niha* 'person' to form the longer expression *samesa niha* 'one person.' However, most numeral expressions for counting people entail the prefix *da-*, or *ra-* in the case of Haloban, attached to the numeral. Ariani (2016) presents *raitu* 'seven people,' whereas Aziz's data presents *daitu* in Devayan. In Sigulai and Nias, *da-* becomes a prefix to all numerals from two to ten, but in Devayan and Leukon *da-* attaches only to numerals from two to seven. In Amery's Haloban data, the prefix *ra-* mirrors the distribution of *da-* in Devayan and Leukon, attaching itself to numerals from two to seven, but this prefix *ra-* does not appear in the data collected by Aziz, but rather, the numerals from one to four combine with *-ata* 'person.' As such, the means by which people are enumerated in these languages is quite distinct from the enumeration of other entities, where the only prefix present is *sa-* 'one,' which corresponds to *sa-* in Jamee and *se-* in Bahasa Indonesia.

### *Indigenous vs Introduced Numerals*

Alongside of the indigenous numerals, the use of introduced numerals is the most obvious source of variation. Recently introduced numerals are borrowed mostly from Jamee and occasionally from Bahasa Indonesia, and are more likely to appear when the number is small than when it is larger. Unsurprisingly, introduced numerals are most likely to appear in combination with introduced measure words, such as *kilo*. Indigenous numerals most likely appear in counting people, animals, and fruits, but considerable variation appears at the point where introduced numerals cut in. Here, Sigulai appears to be the most conservative, as the only one of the four languages indigenous to Pulau Simeulue and Pulau Banyak that uses indigenous numerals above eight. Jamee has most influenced Haloban, as the only language of the four that employs introduced numerals when just counting independently of a specific entity. The other three languages use indigenous numerals at least up to seven for this purpose.

In Ariani's (2016) Devayan data, a similar pattern emerges. Borrowed (Jamee) numerals intrude at several points depending on the object being counted: They appear from two onwards if only counting, from three or four onwards if counting days, but from eight onwards if counting people, animals, or fruits. However, if plants are counted, the Jamee/Indonesian word *tujuh* 'seven' appears rather than the indigenous *itu-* 'seven,' which is used with fruit. The enumeration of weeks, months, or years, borrow all numerals from Jamee. However, some differences between Ariani's (2016) Devayan data and Aziz's 2018-19 Devayan data arise.

#### Difference 1 (Devayan)

Ariani has *itu ngasiha* 'seven animals,' with the indigenous numeral *itu* 'seven,' whereas Aziz has *tujuhngasiha* 'seven animals' with the introduced numeral *tujuh* 'seven.'

### Difference 2 (Devayan)

Ariani has *rufo* ‘two fruits’ compared with Aziz’s *dufo* ‘two fruits.’

In every case for numerals above ten, introduced numbers appear before the teen element. This is also the case further afield in Bahasa Gayo, spoken in the interior of Aceh province. Whilst indigenous numbers appear for one to nine, and are obviously cognate with the indigenous numbers in use in Pulau Simeulue and Pulau Banyak, forms derived from the Malay or Indonesian languages always appear with *belas* ‘teen’ (Anon 2013). Eades (2005: 86) observes that the native Gayo numerals appear for small numbers of animals, people, and items associated with traditional culture, whilst the Malay/Bahasa Indonesia system appears for the introduced measurement of time, weight, length, and prices on goods. As such, Gayo closely parallels this aspect of the counting system of the Barrier Islands’ languages.

### *The Ligature –nga*

An epenthetic syllable *–nga* often appears throughout the data when enumerating certain objects. Its appearance becomes most evident in Devayan and is least apparent in Haloban, and occurs more often in larger numeral expressions. In Devayan, Sigulai, and Leukon, *–nga* appears with all numerals when counting animals, but never in counting thin, flat objects. For example, in its distribution in counting long cylindrical shapes, the following occur: In Nias it occurs with all numerals except one, in Sigulai with three and above, in Leukon with four and above, in Devayan with six and above, and in Haloban just with one, two and five. By contrast, in counting bundles of firewood, it only occurs in Devayan and then only with numerals three and above. In counting fruits, it only appears in Nias, and then just with numerals three and above.

This ligature or epenthetic *–nga* seems to have its origins deep in the DNA of Austronesian languages. It appears in Old Javanese where an epenthetic *–ng* is affixed to the numeral when followed by *atus* ‘hundred,’ *iwu* ‘thousand’ and *yuta* ‘million,’ but not when followed by *welas* ‘teen’ (Miradayanti n.d.: 2, 5-6). Blust (2012) discusses the multiplicative ligature *\*ŋa* in Proto-Malayo-Polynesian which has reflexes in many languages, including Devayan (referred to as Simalur in Blust).

### *Minor Variation in the Form of the Numeral*

All the languages of Pulau Simeulue and Pulau Banyak exhibit minor variation in the form of their numerals, irrespective of whether these numerals are indigenous or introduced, a system predicated on the object being counted. This variation is akin to ‘doubleting,’ as described by Blust (2013: 338-340), except that in this case, we see more than two variants and sometimes up to four or five variants of the one root, with three variants being common within the same language. The indigenous morpheme for ‘two’ sometimes appears as *ro*, sometimes *du*, and sometimes *do*, whilst the introduced

morpheme for ‘two’ varies between *dua* and *duo*. Indigenous elements for ‘three’ vary between *telu*, *telo*, and *tlo*, whilst the introduced ‘three’ sometimes appears as *tigo*, and sometimes as *tiga*. The indigenous element for ‘four’ varies between *hat*, *ata*, *aat*, and *aka* in Devayan, Leukon, and Haloban, compared with *efa* and *eva* in Sigulai, whilst the borrowed numbers are *ampek* and *ampat*. The element for ‘six’ varyingly appears as *nemma*, *nəmma*, *neme*, *nama*, *ennem*, *annam*, *enəm*, *enam*, *ene*, *ena*, and *nem*, as minor variants with a core nV, where V is either e, ə, or a. An additional vowel (e or a) can precede n, whilst m or m plus a vowel (either e or a) can follow the core syllable. We might represent the structure of the variants for ‘six’ as (V1)n(n)V1(m)(m)(V2) where V1 = e, ə or a; V2 = e or a. However, the base form for ‘six’ is never more than two syllables. If the initial V1 is present, the final V2 is absent. The cases of ‘five’ and ‘seven’ show less variation. Only the two variants *lima* and *limo* exist for ‘five,’ as the Indigenous and non-Indigenous numeral coincides. The Indigenous numeral ‘seven’ varies only between *itu* and *fitu*, while the non-Indigenous numeral form is either *tuju* or *tujuh*.

Variation in the base numeral form is achieved through the alternation of d and r; v and f; p and f; k and t; a and o; u and o; a, e and ə; and the addition or deletion of vowels and occasionally syllables, and infrequently the gemination or lengthening of consonants or vowels. We found other instances of doubling in Sigulai: *dadaw* ‘to sit on the floor’ vs *tataw* ‘to sit raised on a chair’; *bala* ‘compact red’ vs *afala* ‘dispersed red’ (see Amery 2016). These variants, in both form and meaning, are obviously the result of a minor modification of the original form, and are similar to examples cited by Blust (2013: 338), such as \**Ratas* > *ratah* ‘human breast milk’ next to *gatas* ‘store-bought milk’ in Tiruray of the Southern Philippines.

Kähler (1961), in his investigation of Simeulue more than half a century ago, reports similar variation to that found in our investigations of numerals with *daro*, *do*, *dua*, *duo*, *ru* and *rua* all glossed as ‘two,’ *təlo*, *təlu* and *tigo* ‘three’; *at*, *aka*, *ampe?* and *dahad* ‘four’; *lima* and *limo* ‘five’; *anam* and *nəm* ‘six.’ Presumably, Kähler’s *daro* and *dahad* have utility when counting people, corresponding to *daro* and *dahat* respectively, which we recorded for the counting of people in Devayan and Leukon.

## The Meaning of Variation

It seems that these languages are hungry for variant forms, and that there is clearly a need for more in-depth investigation of this phenomenon. We obtained our data primarily through elicitation, and built our data on the knowledge of a very small number of speakers of each of these languages. But it is clear that not all speakers agree all the time. We even found the same speaker to volunteer different numerals on different occasions for the same purpose. Significant differences are evident in Amery’s Haloban data, compared with that collected by Aziz, and there are clear differences between Ariani’s (2016) Devayan data and that collected by Aziz. There are also signs of internal inconsistency within data collected by the same researcher. At this stage, it is unknown

to what extent this variation has significant social meaning. A range of sociolinguistic variables, including age, sex, education, and occupation of the speaker should be investigated. Perhaps there is localized subdialectal variation from one village to another. Language shift is likely to be more advanced in some locations than others. The Haloban data would seem to support this hypothesis.

## Conclusion

This investigation has presented an especially interesting case in the study of number systems and counting phenomena, where the extent of variation in the form of the numeral is perhaps unrivalled. In this paper, we have established the origins of this variation as predicated on multiple sources; language contact with Jamee and Indonesian; long-standing variation within Austronesian languages, which has been preserved in this set of languages; fusion between the numeral and the ‘fruit’ classifier; and a process akin to doubling involving minor sound changes and modification of the numeral roots themselves. These processes have resulted in up to five variants of the base numeral form within the one language, even prior to the addition of the semi-productive classifiers and the ligature *-nga*.

Our study constitutes an initial exploration of a complex topic that warrants further investigation through the gathering of spontaneous data from a wider range of speakers before these languages cease to be spoken fluently. This is particularly urgent in the case of Haloban. In a scenario of language attrition, intricate systems of enumeration, such as these, are likely to be amongst the first casualties.

## References

- Aikhenvald, A.Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.
- Alwi, H., Dartjowidjojo, S., Lapoliwa, H., and Moeliono, A.M. (2003). *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Pusat Bahasa dan Balai Pustaka.
- Amery, R. (2016). *Research Report: A Survey of the Languages of Pulau Simeulue and Pulau Banyak*. Submitted to RISTEKDIKTI and CoEDL.
- Amery, R., and Aziz, Z. A. (2019). Enumeration and Classifiers in Pulau Simeulue/Pulau Banyak Languages, Aceh. *SOAS GLOCAL CALA 2019 Proceedings*. [https://glocal.soas.ac.uk/cala2019-proceedings/#tab\\_full-proceedings](https://glocal.soas.ac.uk/cala2019-proceedings/#tab_full-proceedings)
- Anonymous. (2013). *Bilangan Dalam Bahasa Gayo*. [Numbers in the Gayo language] (<http://hericibro.blogspot.com/2013/05/bilangan-dalam-bahasa-gayo.html> Viewed 16 April 2019)
- Ariani, F. (2016). *Numeralia Dalam Bahasa Devayan*. (Unpublished undergraduate thesis). Banda Aceh: Syiah Kuala University.
- Aziz, Z. A., and Amery, R. (2016). *A Survey on the Status of the Local Languages of Pulau Simeulue and Pulau Banyak and their Use Within the Community*.

- Proceedings of English Education International Conference*, [S.l.], v. 1, n. 2, p. 487-490, Nov. 2016. ISSN 2527-8037. Available at:  
[<http://capeu.unsyiah.ac.id/proceedings/index.php/eeic/article/view/95>]
- Blust, R. (2009). *The Austronesian Languages*. Canberra: Pacific Linguistics.
- Blust, R. (2012). The Proto-Malayo-Polynesian Multiplicative Ligature \*ŋa: A reply to Reid. *Oceanic Linguistics*. 51(2): 538-566.  
[<http://www.jstor.com/stable/23321867>]
- Blust, R. (2013). *The Austronesian Languages. Asia-Pacific Linguistics Open Access Monographs*. Canberra: Australian National University.
- Chaer, A. (2006). *Tata Bahasa Praktis Bahasa Indonesia*. Jakarta: Rineka Cipta.
- Chaer, A. (2008). *Morfologi Bahasa Indonesia (Pendekatan Proses)*. Jakarta: Rineka Cipta.
- Chaer, A. (2012). *Linguistik Umum*. Yogyakarta: Rineka Cipta.
- Comrie, B. (2013). Numeral Bases. In Matthew S. Dryer, and Martin Haspelmath (eds.). *World Atlas of Language Structures Online*. Munich: Max Planck Digital Library.
- Daly, P. (2015). *The Aftermath of Aid: Assessing Safety and Sustainability in Aceh, Indonesia a Decade after the 2004 Tsunami*. Annual International Conference (AIC), Universitas Syiah Kuala, Aceh.
- Dixon, R. M. W. (1982). *Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax*. Berlin: Mouton.
- Dryer, M. S. (1992). The Greenbergian Word Order Correlations. *Language* 68: 81-138.
- Durie, M. (1985). *A Grammar of Acehnese: On the Basis of a Dialect of North Aceh*. Dordrecht: Foris.
- Eades, D. (2005). *Grammar of Gayo: A Language of Aceh, Sumatra*. Canberra: Pacific Linguistics
- Gil, D. (2013). Numeral Classifiers. In: Dryer, Matthew S. and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.  
[ <http://wals.info/chapter/55>] (Accessed on 2019-04-21)
- Greenberg, J. H. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Greenberg, Joseph H. (ed.), *Universals of Human Language*, 73-113. Cambridge, Mass: MIT Press.
- Greenberg, J. H. (1972). Numeral Classifiers and Substantival Number: Problems in the Genesis of a Linguistic Type. *Working Papers on Language Universals*, 9: 1-3
- Hale, S., and Napier, J. (2013). *Research Methods in Interpreting: A Practical Resource*. London: Bloomsbury Academic.
- Kähler, H. (1961). *Simalur-German, German-Simalur Dictionary* [Simalur-deutsches Wörterbuch mit deutsch-simaluresischem Wörterverzeichnis]. Berlin: D. Reimer.  
[<https://abvd.shh.mpg.de/austronesian/language.php?id=994>]
- Kridalaksana, H. (2007). *Kelas kata Dalam Bahasa Indonesia*. Jakarta: Gramedia.
- Khaliffitriansyah. (2002). Numeralia Bahasa Kairi Dialek Torai. *Jurnal Kebahasaan dan Kesastraan*, 2(6): 45-64

- Klamer, M. and Kratochvil, F. (2014). The Expression of Number in Languages of East Nusantara: An Overview. In Klamer, Marian and Kratochvil, František (eds.) *Number and Quantity in East Nusantara*. Canberra: Asia-Pacific Linguistics, pp. 1-14.
- Marnita, R. (1996). *Classifiers in Minangkabau*. MA thesis, Australian National University: Canberra.
- Miles, M. B., Huberman, A. M., and Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (Third edition). Thousand Oaks: SAGE Publications.
- Miradayanti, D. A. C. (n.d.) *Numeralia Bahasa Jawa Kuno* [Old Javanese Numerals] *Sastra Jawa Kuno Fakultas Sastra Universtas Udayana* [Old Javanese Literature, Literature Faculty, Udayana University] <https://ojs.unud.ac.id/index.php/sastra/article/view/6573/5040/>
- Newnham, J. and Wackett, E. (2006). *Bagus Sekali!* Port Melbourne, Vic.: CIS·Heinemann.
- Nurjanah, Hiroki, I., and Watanave, H. (2017). Aceh paleotsunami archive visualization for sustainable disaster risk reduction and global information. *International Journal of Asia Digital Art and Design* 21(3): 107-116.
- Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Pearson Longman.
- Retnaningsih, N.S. (2015). *Analisis Kontrastif Numeralia Dalam Bahasa Indonesia dan Bahasa Inggris dan Implikasinya Terhadap Pembelajaran Bahasa Inggris Untuk Penutur Asli Bahasa Indonesia*. Gajah Mada University. Retrieved March 2, 2018, from [[http://etd.repository.ugm.ac.id/index.php?act=viewandbuku\\_id=84580andmod=penelitian\\_detailandsub=PenelitianDetailandtyp=html](http://etd.repository.ugm.ac.id/index.php?act=viewandbuku_id=84580andmod=penelitian_detailandsub=PenelitianDetailandtyp=html)]
- Schapper, A., and Hammarström, H. (2013). Innovative Numerals in Malayo-Polynesian Languages outside of Oceania. *Oceanic Linguistics* 52(2): 423-456.
- Schapper, A., and Klamer, M. (2014). Numeral systems in the Alor-Pantar languages. In Marian Klamer, (ed.) *The Alor-Pantar Languages: History and Typology*. Berlin: Language Science Press.
- Sneddon, J.N., Adelaar, A., Djenar, D.N., and Ewing, M.C. (2010). *Indonesian Reference Grammar*. 2<sup>nd</sup> Edition. Crows Nest, NSW: Allen and Unwin.
- Suliyanti. (2013). *Perbandingan bahasa Jamee dialek Singkel dengan bahasa Jamee dialek Sinabang*. (Unpublished undergraduate thesis). Syiah Kuala University: Banda Aceh, Aceh.
- Tim Balai Bahasa Aceh (2012). *Inilah Bahasa-Bahasa di Aceh* [These are the languages in Aceh] Banda Aceh: Balai Bahasa.
- Woods, D. L. (2011). Counting and Marking Time from the Precolonial to the Contemporary Tagalog World. *Philippine Studies*, 59(3): 337-365.
- Zulfadli. (2014). *A Sociolinguistic Investigation of Acehnese with a Focus on West Acehnese: A Stigmatised Dialect*. PhD Thesis. The University of Adelaide.

## Acknowledgements

The authors thank the anonymous reviewers provided by the JALA for their detailed comments on an earlier version which helped to shape this paper into its present form. Most importantly, we express our sincere thanks to the speakers of Sigulai, Leukon, Devayan, Haloban and Nias with whom we worked.

---

## Endnotes

- <sup>1</sup> Hans Kähler (1961) Simeulue Vocabulary  
<https://abvd.shh.mpg.de/austronesian/language.php?id=994>. (Accessed 15/1/2019)
- <sup>2</sup> *Sara* is the numeral 'one' in Gayo (Eades, 2005: 86) and many other languages of Sumatra (Rosenfelder)
- <sup>3</sup> Numbers from 1 to 10 in over 5000 languages compiled by Mark Rosenfelder  
<http://www.zompist.com/numbers.shtml>. (Accessed 15/1/2019)