Parole: Journal of Linguistics and Education, 14(2), [2024], [43-55]
Available online at: http://ejournal.undip.ac.id/index.php/parole

The Acehnese Numerals in the Aceh Jaya District

Septhia Irnanda [a*], Ida Muliawati [b], Yulsafli Yulsafli [c], Kismullah Abdul Muthalib [d], Nurul Inayah [e]

[ac]*Universitas Serambi Mekkah, Aceh, Indonesia*
[b]*Universitas Iskandar Muda, Aceh, Indonesia*
[de]*Universitas Syiah Kuala, Banda Aceh, Indonesia*

A B S T R A C T

This study examines the phonological varieties of Acehnese numeral words, using primary data from ten villages in the Aceh Jaya district on the west coast of Aceh, Indonesia, and supplementing it with secondary data from Standard Acehnese, Proto-Chamic, and other relevant languages and proto-languages. The research aims to elucidate the historical relationship between the Acehnese and Chamic languages through a comparative analysis. Fifteen numeral words (one to ten, eleven, twelve, twenty, one hundred, and one thousand) were elicited from informants aged 40 to 65. The collected data were recorded, transcribed into IPA symbols, and subsequently analysed. The findings indicate that the split between the Acehnese-Chamic branch and the Malayic languages occurred during the period when the terms for 'eight' and 'nine' were still undergoing lexical stabilisation. The Acehnese language was then subgrouped from Chamic, initially as a dialect. A few centuries later, the ancestors of the Acehnese migrated from mainland Southeast Asia to Sumatra, where their numeral system was subsequently influenced by borrowing from the Malay language. The dialects spoken on the west coast of Aceh retain linguistic evidence of the presence of a language similar to Proto-Chamic and its early contact with indigenous Sumatran populations, and those from the Malay Peninsula.

A R T I C L E I N F O

Paper type:
Research Article

Article history:
Received: 09/01/2025
Revised: 14/07/2025
Accepted: 28/08/2025

Keywords:
- Numerals
- Acehnese
- Historical Linguistics
- Chamic
- Sumatra

1. Introduction

The Austronesian people are believed to have originated from Taiwan around 5,000 years ago (R. Blust, 1995). Their language serves as the ancestor of various languages spoken throughout Southeast Asia, numerous islands in the Pacific, and even some languages in Madagascar, Africa. Outside of Taiwan, Austronesian languages are classified as Malayo-Polynesian languages, which include groups such as Malayic, Acehnese, Javanese, Madurese, Sundanese, Bataknese, and others found in the maritime regions of Southeast Asia.

In mainland Southeast Asia, Austronesian languages are present, but there are fewer speakers. Notably, in countries like Cambodia, Vietnam, Laos, and Thailand, there are minority groups that speak Chamic languages, a specific subset of the Austronesian language family that has established its presence on the mainland (Brunelle et al., 2007; Grant & Sidwell, 2005). The Acehnese language of Sumatra, although surrounded by Malayic languages and other Austronesian languages such as Bataknese, Gayonese, and minorities of Javanese variants, is furtherly related to those languages, and closely related to the Chamic languages of the mainland. It is still a debate whether Acehnese was one of the Chamic dialects in the past, or a sister of it. Thurgood (1999) claims that Acehnese is Chamic, thus placing Acehnese under the Proto-Chamic (1999). However, researchers like Blench (2010) proposed Acehnese to be a separate group that never migrated to the mainland from both the Malayo and Chamic languages' original place, the Borneo of SEA Islands. Dyen (1971) proposed a scenario in which the mainland Chamic languages originated from Sumatra, opposite to Thurgood's and other linguists such as Durie (1985), Cowan (1991), and Sidwell (2005, 2006a). Recently, Irnanda (2024a) proposed that Acehnese is one of the Chamic dialects spoken on the mainland, whose speakers

* septhia.irnanda@serambimekkah.ac.id (Septhia Irnanda)

around the 8th century CE moved to Sumatra after Srivijaya's series of attacks launched on the mainland (Suzuki, 2019).

Therefore, the status of Acehnese within the Malayo-Polynesian language sub-family is complicated due to its proximity and prolonged contact with other closely related Malayo-Polynesian languages, particularly within the Malayic group, which is a sister language to Chamic. The Acehnese lexicon contains a significant number of Malayic borrowings, making it easy to confuse certain forms with the basic vocabulary derived from Chamic Austronesian sources (Thurgood, 1999, 2007). Additionally, the Chamic languages on the mainland have maintained contact, primarily through trade in and around the Gulf of Thailand, with their Malayic counterparts (Coedes, 1975). Champa and Java inscriptions found in Vietnam and Java islands mention the relationship of the two regions in the past (Effendy, 2021). To gain a deeper understanding of the linguistic evolution of Acehnese and its modern form as we encounter it today in the northern tip of Sumatra, further studies on Acehnese dialects are essential.

Therefore, by comparing the innovations and retentions of numerals in the Acehnese and Chamic languages, this study aims to determine the relationship of the two language groups that split in the past.

Numerals play a crucial role in historical linguistics and comparative study. It is one of the basic vocabulary domains that is highly stable, thus is reconstructible (Campbell, 2021; Crowley & Bowern, 2010). The evidence of numeral stability is proven in the reconstruction of the Proto-Indo-European language (Mallory & Adams, 2006). Furthermore, the Sino-Tibetan family, which includes Chinese, Tibetan, Burmese, and many other languages, also shows remarkable stability in its basic numerals, allowing for reconstruction to Proto-Sino-Tibetan (Matisoff, 2003).

Theoretically, the lower numbers like 1-5 are usually more stable and rarely lexically placed compared to numbers such as 7-9 (Campbell, 2021). Often, in many language families, numbers like 6-9 are more susceptible to being borrowed, especially when cultures come into contact with more dominant or economically advanced societies that have more robust or efficient numeral systems. The Kanashi language, for instance, has borrowed most of its numerals from 6 upwards from Indo-Aryan, retaining only 1-5 and one term for 20 from its original Tibeto-Burman stock (Saxena & Borin, 2022). Malay language -teen cardinal, *belas*, is borrowed from Old Javanese (Tudmor, 2007).

One of the arguments for grouping Acehnese, Chamic and Malayic languages into one group, the Malayo-Chamic by Blust (R. Blust, 1994), is based on the comparison of the numerals. Thus, the Malayo-Chamic societies inherited the roots for numbers one to six from the PAN, yet developed independent innovation in 'seven' etyma, from the PAN **siwa* to PMP **tuzuq* (See Table 1). This new word was proposed to have developed from the root 'index finger' or 'to point', or what is known in Modern Malay as *tunjuk* (R. Blust, 1994). In historical linguistics, the function of the hand fingers in the development of numeral roots is something phenomenally encountered in many languages. For example, in the Ali language (Central Africa), "five" and "ten" are respectively *moro* and *mbouna*: *moro* is actually the word for "hand" and *mbouna* is a contraction of *moro* ("five") and *bouna*, meaning "two" (thus "ten" = "two hands")(Ifrah, 2000).

Acehnese and Chamic grouping by Thurgood (1999) is also based on this innovation on 1-7 roots. However, Thurgood (1999) neglects the discrepancies of 8 and 9, the 'teen' and 'ty', in Acehnese and all Chamic languages, which actually can be an important clue to the Acehnese-Chamic linguistic and historical relationship. Therefore, in the present study, the roots of 8, 9 and numeral systems like 'teen', 'ty', 'hundred' and 'thousand' in Acehnese and Chamic are compared to each other and to other relevant languages like Malay and Javanese, to determine the approximate date of the Acehnese split from Chamic.

Furthermore, the numerals across the Acehnese dialects show universality in lexical forms, but high varieties in the phonological level. The consistent lexical forms indicate a shared proto-form and a firm formation of the numeral system before the breakup of the language into dialects. Meanwhile, the variations on the phonological level could also indicate interaction with other language speakers; thus, the Acehnese words change their sounds due to contact with different language speakers, which ultimately results in dialects.

From the four major dialects proposed in Asyik (1987), most western sub-lects of Great Aceh, or modern Aceh Jaya district, demonstrate unique phonological characteristics in their numeral items. Similar to the Great Aceh, some sub-lects in the Aceh Jaya areas also demonstrate peculiar pronunciations, including in their numeral systems. Therefore, the present study aims to describe and explain the phonological peculiarities of the Aceh Jaya dialects to understand their position in the Acehnese dialects grouping, and finally to see the implications of this understanding for the study of Acehnese relationships with Chamic.

The research questions of the present study are as follows:

1) Using the comparative method, how different are the Acehnese and Chamic in numeral roots and systems? What does the result mean in terms of the Acehnese-Chamic relationship?
2) Using the data from the numerals of the Aceh Jaya and Standard Acehnese dialects, how close is Acehnese to Proto-Chamic?

2. Methods

Two main sources of data are employed in the analysis. The first one is the field data collected from nine localities within the district of Aceh Jaya from March to July 2024. The other one is the secondary data gathered from previous studies on Austronesian languages. The Chamic language data were from Thurgood (1999), the North Acehnese dialect was from Durie (1985) and Daud and Durie (1999), and the Proto-Malayic data were from Adelaar (1992). Sources of other languages, including Old Javanese, Austroasiatic, Sanskrit, Tamil, and other relevant ones, are cited whenever they are mentioned in the analysis and discussion.

Primary Data

A fieldwork was carried out using an elicitation technique involving nine Acehnese speakers in Aceh Jaya district from March to July 2024. The nine localities were chosen using snowball and network techniques; the researchers did preliminary research to find out the villages known for their peculiar Acehnese dialects. After that, an informant was chosen from every village using purposive sampling; the person had to be a local, a minimum age of 40, who had never lived in a different place outside of their village. There were 10 informants from 10 villages involved in total. Their age, gender and village name are listed below:

1. A 58 year old female from Sabet
2. A 43 year old female from Gle Jong
3. A 60 year old female from Teumareum
4. A 64 year old female from Seumantok
5. A 48 year old male from Fajar
6. A 67 year old male from Lhok Bot
7. A 50 year old female from Paya Seumantok
8. A 70 year old female from Pulo Ie
9. A 55 year old female from Tuwie Priya
10. A 45 year old male from Alue Punti

Secondary Data

The etymological dictionaries of various sources were consulted, yet mainly the researchers used:

1. Daud & Durie (1999) for Acehnese
2. Thurgood (1999) for Proto-Chamic (PC)
3. Lafont & n-văn-Trọng (1968) for Modern Jarai
4. Adelaar (1994) for Proto-Malayic (PM)
5. Zoetmulder, P.J. (1982) for Old Javanese
6. ACD website for PAN and PMP (R. Blust et al., 2023)

Data Analyses

The data analysis follows the principles of linguistic comparative studies. To answer RQ1, the steps of data analysis are as follows:

1. The forms of numerals were carefully examined across Acehnese and Chamic to determine whether they were descended from the same proto-form or had been borrowed.
2. The phonological forms were compared to see the regular sound correspondences. If the sound pattern was regular, then the words are reflexes from the same proto-form.
3. The morphological forms were then examined to support the shared ancestry. Especially for roots of seven, eight, nine, ten, teen, ty, hundred and thousand.
4. The forms from other related languages, starting from semantic, lexical, phonological, and morphological forms, were checked to confirm whether a certain set has an ancestral relationship or not.

For RQ2, the data analysis also followed the comparative studies. Yet, as the lexical and morphological forms were identical across dialects, the focus was on the phonological patterns and their resemblance to PC as the oldest form. Thus, the oldest dialect can be reconstructed.

3. Results and Discussion

3.1 Acehnese & Chamic Numeral Roots and Systems

This part discusses the results related to RQ 1 of the study. Table 1 compares numerals in Acehnese with Proto-Chamic and some other related languages. This was done to see if the Acehnese numerals were inherited from the same proto-form of Chamic, which is represented by Thurgood's (1999). As suggested in Blust & Smith (2014), both languages have been reportedly descended from Proto-Malayo-Polynesian (PMP), a language that split from the Proto-Austronesian (PAN) language around 3000 years BP around the islands of the Philippines. .(R. A. Blust & Smith, 2014). The morphological forms of some numerals in Acehnese resemble those in Malay, but it remains unclear whether they were borrowed from Malay after the speakers migrated to the SEA islands or inherited from the PC. Therefore, it is important to compare the datasets with some other related languages. Malayic, whose Proto-form is reconstructed by Adeelar (1994), is the cousin language of Chamic and Acehnese; therefore, the sets from Proto-Malay (PM) are included in the comparison. The same thing applies to the Javanese language, which historically is reported to have been a dominant language along with Malay since the Srivijaya era or even earlier (Coedes, 1975). Furthermore, Jarai was chosen to represent the data of modern Chamic languages. Please note that additional data from other relevant languages, such as some other modern Chamic languages, are not listed in the Table, but could be included in the analysis whenever necessary.

Table 1. Numbers in Acehnese and Other Languages

Gloss	**PAN** **Blust (1999)**	**PMP** **Blust (1993)**	**Proto Malayic** **Adelaar (1992)**	**Acehnese (Daud & Durie, 1999)**	**Proto Chamic** **Thurgood (1999)**	**Modern Jarai** **Lafont & n-văn-Trọng (1968**	**Old Javanese** **Zoetmulder, P.J. (1982) (Ngoko)**
One	*esa	*esa	*əsaʔ	sa	*sa	sa, ha	siji
Two	*duSa	*duha	*dua(ʔ)	dua	*dua	dua	loro
Three	*telu	*telu	*təlu	lhɛə, lhɛ:	*klɔw	klâo	telu
Four	*Sepat	*epat	*əmpat	pɯət, pɯt	*pa:t	pă	papat
Five	*lima	*lima	*limaʔ	limʌŋ, limənŋ, limɔŋ	*lima	rɤma	lima
Six	*enem	-	*ənəm	nam	*nam	năm	nem
Seven	*pitu	-	*tujuh	tujoh	*tujuh	tɤjuh	pitu
Eight	*walu	-	*dua(ʔ) alap-an/ *(ə)saʔ alap-an	lapan	*dua-lapan	sapăn, sɤpăn	wolu
Nine	*Siwa	-	*əsaʔ ambil-an	si-kureueng	*samilan *sa-lapan	duapăn	sanga
Ten	*sa-puluq	-	*sA-puluh	si-ploh	*pluh	pluh	
Eleven		-	-	si-blah	-	pluh-sa	
Twenty		-	*dua(ʔ) puluh	dua-ploh	*dua plɔh	duapluh	rwa-ŋ-puluh
fifty				lim(ʌ/ɔ/ə)ŋ-ploh		rɤma pluh	lima-ŋ puluh

To begin the comparison, one should look at the first five roots, 1-5, in Acehnese. Compared to PC, the proto-language suspected to be its ancestor, Acehnese and Chamic both lost the PMP glottal initial syllable of

the root 'one', which must have been due to the Austroasiatic penultimate syllable influence. For the root 'three' Acehnese *lh-* initial could have evolved from the PMP **t-l-,* which shifted to **kl-* in Thurgood's (1999) PC. Given this was these circumstances, the Acehnese split from other Chamic languages soon after the Acehnese-Chamic was developed. Acehnese, with its *tl-* initial for 'three', could have been spoken in the southern part of the mainland, whilst the *kl-* variant should have been spoken on the northern coast of today's Vietnam, which is evidently by the Tsat language that reflects the gloss as *kiu*^{33}(Thurgood, 1999).

The adoption of *tiga* from Sanskrit for 'three' observed in many Malayic languages did not affect either Acehnese or Chamic. We speculate that some Malayic languages had adopted this Sanskrit *tiga* to replace the PAN **telu* sometimes during the Indianisation period. The Malagasy language in Africa, whose speakers were believed to have arrived there from Borneo around 400 AD (Howe, 2021), retained the PMP **telu*, which leads to a speculation that the borrowing of *tiga* into Malayic languages occurred after that. The Old Javanese Krama dialect, or the Javanese high dialect, also has this borrowing, while the Ngoko, the low Javanese dialect, retained the PAN **telu*. The royals of Java could have been the first Austronesian speakers who borrowed the Sanskrit *tiga* into the Austronesian language. Their intermix with Malayic speakers, sometime around the formation of Srivijaya in southern Sumatra, resulted in the Malayic languages borrowing the word *tiga* into their language. Acehnese speakers' encounter with the Malayic speakers in Sumatra should have occurred a few centuries after Srivijaya was established, or after the *welas* '-teen' from the Javanese were absorbed and settled in the Srivijayan Malay language as *belas,* which later borrowed into Acehnese as *blah*.

For the PAN etyma **lima*, Acehnese seems to be the only language among the Malayo-Chamic group that has a final -ŋ reflex; *limʌŋ, limәŋ, limɔŋ*. There are two possible reasons for this phenomenon. Firstly, it is an independent innovation occurring to PAN final -a etyma, such as in *ina>inoŋ* 'mother, woman', and *liŋa>liŋɔŋ* 'sesame' (See section 3.8). Secondly, this final *-ŋ* could also be the result of mutation due to the PMP final **ŋa/*ŋ* linker for multiplying tens, as in Old Javanese, *lima-ŋ puluh* means fifty (Zoetmulder, 1982). It then raises a question of why the mutation only occurred in Acehnese's five *limong*, and did not occur in the other numbers, such as in the Javanese case? One possible explanation is that the Acehnese 'five' final *-ŋ* is not a recent innovation that occurred post-Acehnese dispersal to Sumatra, but an innovation that occurred way earlier on the mainland. This assumption is based on two important evidence; first, a Chamic language—called Cac Gia Roglai- that was reported to have the same velar nasal accretion as Acehnese (Lee, 1998), with highly basic vocabularies such as 'five', 'ricefield', and 'flower'. Secondly, the absence of final *-ŋ* on numbers like 'three' as found in Old Javanese numeral systems is evidence itself that the final *-ŋ* in Acehnese is not a morphological rule adopted from Javanese but purely a phonological one. Since it occurred on the low-number root, the five, the innovation must have occurred quite early.

Whilst the lexical form of the root 'six' is retained, the roots 'seven', 'eight' and 'nine' in the Malayo-Chamic group, including Acehnese were replaced by what Blust (1994) proposed as independent innovations. The lexical innovations of Malayic-Chamic-Acehnese 'seven' from PAN **pitu* marked the split of these three languages from their PMP ancestor, forming an independent group, distinct from other Austronesian languages. This is parallel with Smith's (2005, 2006b; 2017) proposal that challenges Blust's (1994) Western-Malayo-Polynesian language group members, in which he included Javanese, Balinese, Sumbawa, Malay and all Sumatran languages into one group. This 'seven' innovation in both Malay and Chamic hints at a period in Borneo, before the migration of Chamic ancestors to the mainland.

The development of 'eight' and 'nine' seems to have also occurred globally in the three languages of Acehnese, Chamic and Malay, as the roots that formed the lexicons are the same; the roots 'alap' and 'ambil' are both PAN-origin words meaning 'to snatch' (R. Blust, 1994). The inconsistency of the usage of these two morphemes is found in both PM and PC, where in the PM, the 'eight' is reconstructed into two versions: **dua(ʔ) alap-an* and **(ə)saʔ alap-an* (Adelaar, 1994). Whilst Thurgood (1999) reconstructs only **dua-lapan* in PC, even though Jarai demonstrates the sa-lapan variant in its 'eight' sapăn. Acehnese lost its first morpheme, thus its lapan should have been from PC. These discrepancies demonstrate the period of branching among these languages of Malayic and Chamic.

Meanwhile, the gloss 'nine' is reconstructed as only one form, the **əsaʔ ambil-an* in PM, and two forms in PC; **samilan* and **sa-lapan*. Jarai, again, reverse the usage of the first morpheme, thus have *duapăn* for 'nine'. Acehnese peculiarly uses a totally different PAN root for this number. Instead of *alap* or *ambil,* Acehnese employs *kurang*, a PAN term for 'to reduce', which looks like an independent lexical innovation to replace the PC **sa-lapan* as it is phonologically too close to its eight **dua-lapan*. Especially as Acehnese tends to reduce its first syllable of the word, making the 'eight' and 'nine' indistinguishable. Therefore, the Acehnese *si-kuruəŋ*,

we propose, is an innovation to replace the former-lapan root, which it had inherited from PC. Yet, as the root *kuruəŋ* is not a mainland material, this innovation should have occurred after it migrated from the mainland.

Furthermore, the use of *si-* instead of *sa-* in the Acehnese word *si-kuruəŋ* also indicated that this *si-kuruəŋ* innovation occurred when the Acehnese morpheme *sa-* , which means 'one' had been replaced by *si-*. The morpheme *sa-*, which we suggest is the older form of *si-,* according to the Acehnese Dictionary (Daud & Durie, 1999), has two meanings: 'one' and 'same'. Thus, words like *saban, saho, sapat, saboh*, *saneuk*, *sadum,* and *sapeu* subsequently mean 'one destination', 'one place', 'one thing', 'the same amount' and 'none' (for a negative sentence). The morpheme *sa-* was slowly replaced by *si-* due to Malay's *se-* /sə/ influence. Similarly, the prefix *si-* was used as a quantifier for single nouns—usually following the same semantic characteristics of Malay word formation, such as: *sidroe-seorang* 'one person'*, siblah-sebelah* 'one side'*, sikrek-sepotong* 'one piece (for long stick-like object'*, sineuk-sebuah* 'one piece (for round bead-like object)'*, siulah-seulas* 'one piece (for thin leaf-like object)', and many more. *Sa-boh* /sabɔh/ retains in form, while its counterpart, *saneuk* /sanɯʔ/, has almost completely lost it. The shift of the vowel from /a/ to /i/ in Acehnese nine could be the result of the syllabic position. The first and second syllables in Acehnese and Chamic are weak syllables. This characteristic is due to their strong Austroasiatic influence that emphasises the final syllable. Thus, nine in Acehnese could have been *sa-kuruəŋ*, but this form has been lost in all its modern dialects.

The Acehnese language also differs from PC in terms of '-teen' and '-ty' numbers. Acehnese is more like Malay than Jarai, has *ploh* only for -ty, and *blah* for '-teen'. As the Malay 'teen' -*belas* was adopted from Javanese, thus the Acehnese encountered Malay and was influenced by its numerals after the political union of Malay and Javanese speakers, which historically occurred during the rise of the Srivijaya era, around the 7th century CE (Coedes, 1975).

To conclude, based on the result discussed, Chamic and Malayic groups split in Borneo after they replaced the PMP 'seven' *pitu* (R. Blust, 1994). Sometime between 200-300 BC, a group migrated to the mainland and became the Chamic speakers (Brunelle et al., 2007; Grant & Sidwell, 2005; Sidwell, 2005, 2006b).

However, based on evidence of final nasal accretion on root 'five', Acehnese subgrouped from Chamic for the first time not due to Acehnese's ancestors' migration to SEA, but due to the formation of a dialectal chain of PC on the mainland during the first half of the first millennium or sometime around the 5th century CE, as suggested by Sidwell (2005) and also supported by the important and peculiar Austroasiatic vocabularies in Acehnese(Sidwell, 2005, 2006b). At that time of the split, Acehnese was in the middle of developing final velar nasal accretion to its basic words. Sometime, or around the 8th century CE, as archaeological and historical evidence suggested (Irnanda, 2024), Acehnese left the mainland to settle in Sumatra. Then, more phonological changes occurred on its numerals. It borrowed the '-teen' and '-ty' systems from Malay, replaced the PC 'nine' with a new innovative word using the roots 'one-less', copying the same mechanism of PM 'one-taken'.

The section below discusses the numeral phonological discrepancies in Acehnese dialects and how they affect our understanding of Acehnese status as a Chamic language.

3.2 Numerals in Acehnese Dialects: The Aceh Jaya District

This part discusses the results related to RQ 2 of the study. Numerals in the Acehnese dialects are only varied on the level of phonology. Table 2 shows the numerical comparison of the Aceh Jaya dialects and the Standard Acehnese.

Table 2. The Numerals of the Aceh Jaya Dialects from the East to the West

No	GLOSS	SA (North Aceh)	Sabet	Gle Jong	Teum areum	Seum antok	Fajar	Lhok Bot	P. Seum antok	Pulo Ie	Tuwi Priya	Alue Punti
1	*one*	*sa*	*Sə*	*sə*	*Sə*	*Sa*	*Sa*	*Sa*	*sa*	*sa*	*Sa*	*Sa*
2	*Two*	*dua*	*duə*	*duə*	*duə*	*dua*	*dua*	*dua*	*dua*	*dua*	*dua*	*dua*
3	*three*	*lhɛə*	*lheiə*	*lhɛiə*	*lhɛ*	*lɛ*	*lhɛ*	*lhɛ*	*lhɛ*	*lhɛ*	*lhɛi*	*lhɛi*
4	*four*	*puət*	*puət*	*puət*	*put*	*put*	*put*	*put*	*put*	*put*	*put*	*put*

5	*Five*	*limʌŋ*	*limәŋ*	*limәŋ*	*limʌŋ*	*limʌŋ*	*limʌŋ*	*limʌŋ*	*limʌŋ*	*limʌŋ*	*limʌŋ*	*limʌŋ*
6	*six*	*nam*	*nam*	*nam*	*nam*	*nam*	*nam*	*nam*	*nam*	*nam*	*nam*	*nam*
7	*seven*	*tujoh*	*tujoh*	*tujoh*	*tujoh*	*tujoh*	*tujoh*	*tujoh*	*tujoh*	*tujoh*	*tuɟoh*	*tuɟoh*
8	*eight*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*	*lapan*
9	*nine*	*sikuruәŋ*	*SikuRuәŋ*	*SikuRuәŋ*	*SikuRuŋ*	*Sikuruŋ*	*SikuRuŋ*	*Sikuruәŋ*	*sikuRuŋ*	*sikuRuŋ*	*Sikuruәŋ*	*Sikuruәŋ*
10	*ten*	*siploh*	*Siploh*	*Siploh*	*Siploh*	*Siploh*	*Siploh*	*Siploh*	*siploh*	*siploh*	*Siploh*	*Siploh*
11	*eleven*	*siblah*	*Siblah*	*Siblah*	*Siblah*	*Siblah*	*Siblah*	*Siblah*	*siblah*	*siblah*	*Siblah*	*Siblah*
12	*Twelve*	*dua blah*	*duə blah*	*duɔ blah*	*duə blah*	*dua blah*	*dua blah*	*dua blah*	*dua blah*	*dua blah*	*dua blah*	*dua blah*
13	*Twenty*	*dua ploh*	*duə ploh*	*duɔ ploh*	*duə ploh*	*dua ploh*	*dua ploh*	*dua ploh*	*dua ploh*	*dua ploh*	*dua ploh*	*dua ploh*
14	*One hundred*	*sirətoh*	*SәRətoh*	*SiRətoh*	*SәRətoh*	*Surətoh*	*Sərətoh*	*SiRətoh*	*sərətoh*	*sərətoh*	*Sirətoh*	*Sərtoh*
15	*One thousand*	*siribɛ*	*SәRәbei*	*SiRәbeiə*	*SәRәbei*	*SuRәbɛ*	*Sәribɛ*	*SiRibɛ*	*sәRibɛ*	*sәribɛ*	*Siribɛi*	*Sәribɛi*

First, it is essential to understand the phonological system of the Standard Acehnese (SA) dialect, which is the Acehnese dialect spoken in North Aceh regency (Durie, 1985), and how it differs from the dialects in the district of Aceh Jaya.

The first pivotal sound discrepancy between the SA and Aceh Jaya is on the fricative /s/ sound. In SA, /s/ is lamino-dental rather than apical-alveolar. It means that the tongue used to produce this sound is not the tip but the side of the tongue. The tongue tip extends further forward, creating channel turbulence at the teeth (Durie, 1985, p. 15). While in most Aceh Jaya dialects, and also in most Great Aceh district, the /s/ no longer sounds like a fricative, as the turbulence is extremely minimal, the /s/ sound resembles the sound /θ/. Durie (1985) described this non-North Aceh dialect as a lamino-alveodental stop. In this study, we distinguish the two dialectal differences using two symbols: the /s/ for the well-fricated /s/ as in SA, and /S/ for the slightly or minimally fricated /s/ as in Aceh Jaya.

The second important difference is in the trilled sound /r/. In SA, the sound is alveolar, but in Aceh Jaya, the uvular R sound is preferred. In this study, the SA /r/ sound is thus symbolised as /r/, and the uvular sound as /R/.

Regarding the vowels, although the SA and Aceh Jaya dialects have about the same vowel inventories, the positions of each in a word are different. For instance, SA does not allow the schwa sound in a final open syllable, yet Aceh Jaya –although not all dialects- does (see gloss 'one' and 'two'). Aceh Jaya also allows schwa as the nucleus of the weak syllable—usually the first syllable of the word--, while SA either has unrounded /ɯ/ or /i/.

3.2.1 Gloss 'one'

In Aceh Jaya, the initial lamino-dental /S/ is dominant and affects the numeral of 'one'. As illustrated in Figure 2, the lamino-dental sounds /S/, are spread to the whole district from the east to the west. Thus, the Standard Acehese/sa/ whose /s-/ sound is a lamino-dental fricative (Durie, 1985) became a stop in Aceh Jaya.

The presence of these two forms of /s/-/S/ in Acehnese as the initial sound raises a question of whether this fricative-stop lamino sound is the sound of PC before it changes to the fricative apical sound. Durie (1985) described that the Acehnese North Dialect has a fricative lamino-dental /s/ which is realised in Aceh Jaya and Great Aceh dialects as a stop with the same tongue manner. In recent decades, the younger generation who speak less Acehnese and more Indonesian Malay would diminish this laminal position and make the sound apical. Yet, it is clear that the Acehnese laminal fricative/stop sound has been present for a long time, considering its widespread presence in two of the most important Acehnese dialects: the Great Aceh and North Aceh. This sound is also present in the Pidie and West Aceh dialects.

However, is this sound inherited from the Chamic family? Or an independent innovation developed in the SEA island? According to Thurgood (1999), the Written Cham initial /s-/ sound is transcribed as /ts-/. In addition to that, some Malayic basic words, like 'milk' *susu* and 'comb' *sisir* are respectively transcribed as *tathuw*/*tasou*, and *tasi* in Written Cham inscriptions (Thurgood, 1999). Not limited to Written Cham, this

sporadic */S-/* to */t-/* sound shift also occurs in Jarai, Northern Roglai and Phan Rang Cham. Hence, from Thurgood's data of the mainland Chamic languages /s/, the lamino dental fricative initial in Acehnese, we propose, was the sound inherited from Chamic.

The researchers believe that this unique fricative in Acehnese is a retained sound from PC, which is seen from the lamino-dental stops in the Aceh Jaya dialects such as Sabet and Teumareum, or the dialects considered one of the oldest in Durie (1985). Yet, due to the increased popularity of Malay with its apico-alveodental /s/ fricative sound, the laminal /s/ in Acehnese gradually becomes alveo-dental and apical.

Figure 2. gloss 'one'

3.2.2 Gloss 'two'

Moreover, in the eastern dialects, the Lamno and Daya dialects, 'one' and 'two' are reflexed with a schwa vowel ending instead of low /a/ as in the SA dialect. Other than on the east side of Aceh Jaya, this schwa final in open syllables is also common in some dialects in Great Aceh. This final schwa ending is not typical in Sumatra Malayic dialects, like Tamiang in the east side of Aceh Province, or the Minangkabau whose speakers are spread on the south coast of Aceh. Instead, this feature is common in the Malay Peninsula, leading to an assumption that the phonological feature is borrowed from the Peninsular Malay. The lexical comparative study of the Aslian and Acehnese (Irnanda, 2024b) captures an impression of contact between the Acehnese speakers of Sumatra and the Aslian-Malay bilinguals of the Peninsula in old times, probably as early as 500-1000 AD. The availability of final schwa in Acehnese open syllables might be the product of the bilingualism of Chamic-Acehnese and Peninsular Malayic lect in the past.

However, the final schwa is absent in western Aceh Jaya, where the open /a/ is preferred instead (See Figures 2 & 3). In the southern area of the district, the final vowels in gloss 'one' and 'two' are realised as those

in the Standard Acehnese and Proto-Chamic, indicating Aceh Jaya southern dialects' closer genetic linguistic affiliation with the northeast coast dialects.

3.2.3 Gloss 'three'

Acehnese spoken in Jaya, specifically the speakers of the Lamno dialect (village of Sabet in Table 1), realise the diphthong of the open syllable in the root 'three' slightly distinct from the SA. The peculiarity lies in the first vowel of the diphthong; the Lamno people reflect the PC diphthong's back open-mid vowel *-ɔ- as a front close-mid vowel /-e-/. Compare this to the SA's front open-mid /-ɛ-/. See Figure 4.

Figure 4. Gloss 'three'

Based on this finding, it seems logical that the Standard Acehnese reflex of PC *-ɔ- diphthong in the final syllables is an innovation that occurred in the Sumatra area, as no modern Chamic shares this innovation. The front open-mid variants of that in SA should predate the closed-mid ones spoken in Aceh Jaya, as this form is closer to the PC back open-mid vowel. However, a complete comparison using the data of the Pidie and Great Aceh districts should be studied to understand the whole journey of this sound change.

3.2.4 Gloss 'four'

In PC, the nucleus of the one-syllable word 'four' is a long vowel *-*a:*-. Meanwhile, in the SA and the Lamno dialect, it is reflected as diphthong /-ɯə-/, and more to the west of the Aceh Jaya district, it becomes a monophthong */-ɯ-/*. This rise of PC open-front vowel **a:* to the unrounded back close vowel /ɯ/ occurs in many other main closed syllables of PC /**a:/*. According to Irnanda (2024a) this innovation occurred in the mainland and is parallel with the rise of the same PC vowel in Cac Gia Roglai and Phan Rang Cham to unrounded central close vowel */ɨ:/* for glosses of 'betel nut' and 'gold' (Lee, 1998).

The PC long vowel /a:/ is not completely lost in the Acehnese 'four' as both the Standard Acehnese and the easternmost dialects of the Aceh Jaya district retain it in the form of a diphthong. More to the west of the district, the diphthong is reduced to a monophthong. This could have been caused by the employment of Acehnese as a second language by speakers of Batak or other Austronesian natives who dwelt around the region many centuries ago. Alternatively, this monophthongisation could have started recently with the increasing popularity of Indonesian Malay.

Figure 5. Gloss 'four'

3.2.5 Gloss 'five'

Figure 6. Gloss 'five'

As the root 'four', the Acehnese root 'five' has a feature developed in the mainland but is not widely shared with many of today's mainland Chamic languages. Again, as the long vowel in the middle position of PC 'four', which becomes a diphthong in Acehnese, the velar nasal accretion in root 'five' was an innovation that occurred before it reached the island. Therefore, this final nasal accretion is a dialectal uniformity in Acehnese.

However, the 'five' is realised differently by the Acehnese speakers in the Aceh Jaya district in terms of its main syllable vowel. In the eastern area, the vowel is a unrounded central schwa sound, while in the west, the vowel is unrounded back open-mid /-ʌ-/. From the authors' observation, in the West Aceh district that borders Aceh Jaya on the west, the vowel becomes rounded back open-mid /-ɔ-/. Thus, Acehnese rounded and unrounded back open-mid /-ʌ-/ and /-ɔ-/ are minimal pairs in Acehnese. Compared to the Malayic languages around it, Acehnese has a more complex vowel inventory (*Acehnese Phonetics by Asyik*, n.d.; Durie, 1985). The rounded-unroundedness, which is minimal pairs in several Acehnese vowels, might have been a problem for the non-Acehnese speakers. The vowel variations on this feature on gloss 'five' might have been the result of a language shift from other Austronesian languages in Sumatra to Acehnese.

3.2.6 Gloss ‘six’, ‘seven’, ‘eight’, ‘nine’ and ‘ten’

There are not many phonological discrepancies between the SA and the Aceh Jaya dialects’ numerals, ranging from six to ten (See Table 2). The roots of ‘six’, ‘seven’, and ‘eight’ are identical, lexically and phonologically. Root ‘nine’ and ‘ten’ are slightly different; the ‘ten’ is varied in the /s/ sound of the first syllable, while the ‘nine’ is varied not only on the consonant /s-S/ and /r-R/ sounds, but also in the main syllable diphthongs. Some Aceh Jaya dialects do not diphthongise their main syllable in ‘nine’, which might be due to a language shift into Acehnese in the past, probably from one of the Batak-related languages spoken in the interior.

3.2.7 Gloss ‘one hundred’ & ‘one thousand’

Figure 7. Gloss ‘hundred’

Figure 8. Gloss ‘thousand’

The vowel universality in the first syllable of the root ‘nine’, both in SA and Aceh Jaya, is not found in the roots ‘one hundred’ and ‘one thousand’. The ‘hundred’ and ‘thousand’ morphemes are indeed universal across dialects, yet the morpheme ‘one’ preceding them is not (Look at Figures 7 & 8). In isolation, the morpheme is the PC **sa*. Over time, the final vowel shifted into */sə-/,* and then */su-/ and finally /si-/.*

The liquid r[R] is also worth a discussion as it reflects the consonants of Austronesian languages. As seen in Figures 7 and 8, the reflexes with uvular /R/are very dominant in the whole area, except in the western end of the district. As this west end area borders with Pidie district in the north, the alveolar /r/ in their language might be influenced by the Pidie migrants into the area.

Finally, the reflexes of the diphthong sound in the final syllable of the root ‘thousand’ are consistent with those in the root ‘three’; thus, all come from the PC **ɔw*. The open-mid /-ɛə-/ is closer to PC, thus could be proposed as the proto-form or the oldest form of all dialects.

4. Conclusion

The present study provides an analysis based on the comparison of numerals in Acehnese and Chamic, showing that the Acehnese-Chamic-Malayic ancestors diverged while developing their 8-10 numerals. Acehnese and other Chamic languages also separated into dialects not long after the Proto-Chamic formation, as evidenced by the multiple etyma reconstructed for ‘eight’ and ‘nine’. Acehnese ‘nine’ employs a different morpheme from those in the mainland Chamic and Malay languages, which is suspected to be a later lexical innovation that occurred in Sumatra, likely due to the ambiguity of the PC -lapan morpheme, present in both glosses for ‘eight’ and ‘nine’. The PC /s/ laminal sound is retained in both Aceh Jaya and the Standard dialect, whereas in some dialects it has become minimally fricative, resembling the English /θ/ sound. Furthermore, the study highlights

phonological peculiarities, such as final schwa and reduced diphthongs in the Acehnese Jaya dialectal numeral system, serving as evidence of bilingualism between the Acehnese ancestors and the native people of Sumatra who settled the region before the expansion of Acehnese speakers. Future studies should explore more fundamental semantic domains, such as native plants and animals, since numerals alone cannot definitively determine the relationships between languages.

References

Acehnese Phonetics by Asyik. (n.d.).

Adelaar, K. A. (1992). *Proto Malayic:The Reconstruction of its Phonology and Parts of its Lexicon and Morphology*. Pacific Linguistics. https://doi.org/10.15144/PL-C119.cover

Asyik, A. G. (1987). *A contextual grammar of Acehnese sentences*.

Blench, R. (2010). Was there an Austroasiatic presence in Island Southeast Asia prior to the Austronesian expansion? *Bulletin of the Indo-Pacific Prehistory Association*, *30*, 133–144.

Blust, R. (1994). The Austronesian Settlement of Mainland Southeast Asia. In K. A. & T. H. (eds.) (Ed.), *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society 1992* (pp. 25–83). Tempe Arizona, Arizona State University.

Blust, R. (1995). Sibilant assimilation in Formosan languages and the Proto-Austronesian word for'nine': A discourse on method. *Oceanic Linguistics*, 443–453.

Blust, R. A. ., & Smith, A. D. . (2014). *A bibliography of the languages of Borneo (and Madagascar)*. Borneo Research Council, Inc.

Blust, R., Trussel, S., & Smith, A. D. (2023). *CLDF dataset derived from Blust's "Austronesian Comparative Dictionary."* Austronesian Comparative Dictionary. https://doi.org/(v1.2) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7741197

Brunelle, M., Linguistique, D. De, & Ottawa, U. (2007). *Revisiting the expansion of the Chamic language family*. 1–26.

Campbell, L. (2021). *An Introduction*.

Coedes, G. (1975). *The indianized states of Southeast Asia*. University of Hawaii Press.

Cowan, H. K. J. (1991). Achenese Dialects in Connection with Chamic Migration. In R. Harlow (Ed.), *Western Austronesian and Contact Languages*. Linguistic Society of New Zealand Auckland New Zealand.

Crowley, T., & Bowern, C. (2010). *An introduction to historical linguistics*. Oxford University Press.

Daud, B., & Durie, M. (1999). *Acehnese-Indonesian-English Thesaurus*. Pacific Linguistics, Research School of Pacific and Asian Studies, The ….

Durie, M. (1985). *A grammar of Acehnese: on the basis of a dialect of North Aceh*. Foris Publications Holland.

Dyen, I. (1971). The Chamic languages. *Current Trends in Linguistics. Mouton: The Hague*, *8*, 200–210.

Effendy, M. (2021). Champa: Territories and Networks of a Southeast Asian Kingdom, by Arlo Griffiths, Andrew Hardy, Geoff Wade (eds). *Bijdragen Tot de Taal-, Land-En Volkenkunde/Journal of the Humanities and Social Sciences of Southeast Asia*, *177*(1), 144–146.

Grant, A., & Sidwell, P. (2005). Chamic and beyond: studies in mainland Austronesian languages. *Pacific Linguistics*. https://doi.org/10.15144/PL-569.cover

Howe, P. (2021). Central Malagasy. *Journal of the International Phonetic Association*, *51*(1), 103–136. https://doi.org/10.1017/S0025100319000100

Ifrah, G. (2000). *The universal history of numbers*. Harvill London.

Irnanda, S. (2024a). Acehnese: The Dating of Into-Sumatra Migration. *Jol.r*, *3–4*.

Irnanda, S. (2024b). *Aslian Substratum in Acehnese (Was the Contact Direct or Malay-Mediated?)*.

Lee, E. W. (1998). THE CONTRIBUTION OF CAT GIA ROGLAI TO CHAMIC. In D. Thomas (Ed.), *Papers in Southeast Asian Linguistics No. 15: Chamic Studies* (Issue 15, pp. 31–54). Pasific Linguistics, The Australian National University. https://doi.org/10.15144/PL-A89.31

Mallory, J. P., & Adams, D. Q. (2006). *The Oxford introduction to proto-Indo-European and the proto-Indo-European world*. Oxford University Press.

Matisoff, J. A. (2003). *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. Univ of California Press.

Saxena, A., & Borin, L. (2022). *And then there was one: Kanashi numerals from borrowed superdiversity to borrowed uniformity*.

Sidwell, P. (2005). Acehnese and the Aceh-Chamic language family. In *Chamic and beyond: Studies in mainland Austronesian languages.* (pp. 211–245). Pasific Linguistics.

Sidwell, P. (2006a). Dating the separation of Acehnese and Chamic by etymological analysis of the Aceh-Chamic lexicon. *MONKHMER STUDIES*, *36*, 187.

Sidwell, P. (2006b). Dating the separation of AcSidwell, P. (2006). Dating the separation of Acehnese and Chamic by etymological analysis of the Aceh-Chamic lexicon. MonKhmer Studies.ehnese and Chamic by etymological analysis of the Aceh-Chamic lexicon. *MonKhmer Studies*.

Smith, A. D. (2017). The Western Malayo-Polynesian Problem. In *Oceanic Linguistics* (Vol. 56, Issue 2).

Suzuki, T. (2019). *The History of Srivijaya, Angkor and Champa*. Mekon.

Thurgood, G. (1999). From Ancient Cham to Modern Dialects: Two thousands years of language contact and change. In *Oceanic Linguistics Special Publications* (Vol. 28). University of Hawai'i Press. University of Hawaii Press

Thurgood, G. (2007). The historical place of Acehnese : the known and the unknown (A Working Paper). *First International Conference of Aceh and Indian Ocean Studies*.

Tudmor, U. (2007). Borrowed elements in the Malay-Indonesian numeral system. *The Eleventh International Symposium on Malay/Indonesian Linguistics*.