HOKKAIDO UNIVERSITY

Title	Reconstruction of “one” in Proto-Atayalic : In connection with “ten” in some Formosan languages
Author(s)	Ochiai, Izumi
Citation	北方人文研究, 14, 55-70
Issue Date	2021-03-25
Doc URL	https://hdl.handle.net/2115/80886
Type	departmental bulletin paper
File Information	14_05_Ochiai.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

[Article]

Reconstruction of "one" in Proto-Atayalic
In connection with "ten" in some Formosan languages*

Izumi Ochiai
(Center for Ainu and Indigenous Studies, Hokkaido University)

Abstract

This study discusses a newly discovered form of the numeral "one" in the Atayalic subgroup (Atayal and Seediq) of the Austronesian language family. The form of this bound numeral "one" is *xa* (e.g., *taxa* "alone" in Seediq and *saxa* "alone" in Atayal). This study found more than ten words derived from the bound numeral *xa* in Seediq and Atayal. In these words, the bound numeral *xa* accompanies a prefix or both a prefix and a suffix. Various prefixes attach to *xa* in both Seediq and Atayal. As a suffix, only *-l* is used, which is a reflex of *-N "recurrence" in Proto-Austronesian. Based on the Seediq and Atayal forms of this bound numeral "one," the Proto-Atayalic form can be reconstructed as *xa. Furthermore, another Formosan language, Saisiyat, is found to have cognates with these words derived from the bound numeral *xa* in Seediq and/or Atayal. In Saisiyat, the bound form of the numeral "one" is *ha*, which is suggestive of the numeral "one" in Proto-Austronesian, *əsa. This study proposes that the Proto-Atayalic *xa is a reflex of *əsa (or its shortened form *sa), the Proto-Austronesian numeral "one," although the expected reflex in Proto-Austronesian would be *ha according to regular sound correspondences. There are three pieces of evidence. First, only one particular suffix *-N attaches to *xa. This suffix attaches to numerals in Austronesian languages. Second, the reflex of *sa would be *ha* in Saisiyat, and this is the actual morpheme obtained in the cognates of Atayalic languages. Third, some Formosan languages including Seediq have a numeral "ten" originating in a proto-form reconstructed as *ma-sa-N.

Key words: Atayalic, Saisiyat, bound numeral, *xa, *ma-sa-N

1. Introduction

According to Blust (1999), the Atayalic subgroup is a first-order subgroup of the Austronesian

---

* This paper is based on the presentation 'Atayalic numeral one and its word family' given at the 2019 meeting of Symposium for the Research in World Austronesia and Indigenous Peoples, on 14 June 2019, at National Tsing Hua University. A portion of the Seediq data was also discussed in the presentation 'Origin of the numeral 'five' in northwestern Formosan languages: Revision of the Proto- Austronesian numeral system' given at the 25th Annual Meeting of the Southeast Asian Linguistic Society, on 27 May 2015, at Payap University, Chiang Mai. I would like to thank not only those who gave me comments and suggestions on those occasions, but also the reviewers of the earlier version of this paper. However, any errors in this paper are mine.

language family. There are ten first-order subgroups of Austronesian, nine of which belong to the indigenous languages of Taiwan, and these are collectively referred to as the Formosan languages. The other subgroup is Malayo-Polynesian. Formosan languages include Atayal, Seediq, Saisiyat, Bunun, Thao, Pazih, Babuza, Hoanya, Tsou, Saaroa, Kanakanvu, Siraya, Paiwan, Puyuma, Amis, Kavalan, and Basay.

The Atayalic subgroup comprises two languages: Seediq and Atayal. Ogawa and Asai (1935) describe Seediq and Atayal as having each diverged into two dialects. The Seediq dialects are called Paran Seediq and Truku Seediq; the Atayal dialects are called Squliq Atayal and C'uli' Atayal ([ʦɪʔuliʔ]). The sources of data in this paper are as follows. Paran Seediq comes from the notes by the author from fieldwork conducted in the years around 2015. Truku Seediq is taken from a dictionary (Rakaw et al. 2006) unless cited otherwise. Squliq Atayal is taken from another dictionary (Ogawa 1931) unless cited otherwise. C'uli' Atayal is taken from a word list (Council of Indigenous Peoples 2013).

The topic of this paper is the numeral "one" in the Atayalic languages. The forms for that numeral are known to be *kiŋan* in Paran Seediq, *kiŋal* in Truku Seediq, *qutux* in Squliq Atayal, and *ʔutux* in C'uli' Atayal. These numerals are independent forms that can be used alone. For example, *kiŋan* in Paran Seediq means "there is one [thing]." These numerals also precede nouns and noun phrases as modifiers. For example, *kiŋan alaŋ* (one village) in Paran Seediq means "one village." Incidentally, independent forms of the numeral "one" can be reconstructed as *kiŋal in Proto-Seediq and *qutux in Proto-Atayal.

These independent forms for the numeral "one" in the Atayalic languages do not reflect the numeral "one" *əsa of Proto-Austronesian.[1] Ogawa (1944: 469–470) points out that the reflexes of *əsa are retained in *min-xa-l*, which is the Atayal word for "once" and in *ma-xxa-l*, which is the Seediq numeral "ten."[2] Ogawa assumed that the *s in Proto-Austronesian changed to *h* in Atayal and to *x* in Seediq in these cases. In addition, for the Atayal form, Ogawa and Asai (1935: 35) comment that *xa* cannot be used independently. These analyses was basically on point, but it needs detailing.

This paper proposes that there is a bound form of the numeral "one" in the Atayalic languages, as well as independent forms of that numeral. The bound numeral "one" is reconstructed here as *xa in Proto-Atayalic,[3] and *xa is proposed to be a reflex of "one" in Proto-Austronesian, which is reconstructed as *əsa.

---

1) This proto-from is from Blust and Trussel (2010).

2) Ogawa (1944) wrote this Atayal form as *min-ha-l*. However, the present author has modified the medial consonant *h* into *x*. The form for the numeral ten in Seediq, *ma-xxa-l*, is taken from the appendix of Ogawa (1944). The double *x* in his orthograhy seems to indicate the stressed syllable, which is the penlutimate syllable in Seediq.

3) Bound numerals in Formosan languages are presented in Zeitoun, Teng and Ferrell (2010: 868–872). These languages include Rukai, Bunun, Tsou, Saaroa, Kanakanavu, Thao, Saisiyat, Paiwan, and Puyuma. The bound numerals in Atayalic languages were not provided in their paper. Ochiai (2020: 4–5) complemented part of this gap by saying that bound numerals in Paran Seediq are found only in "one" and "two," which are *xa* and *pusa*, respectively.

Among present-day speakers of Atayalic languages, the origin of *xa has been forgotten. It is fossilized in words composed historically by affixation, and the original meaning of “one” is opaque in some cases. These words that historically derive from *xa are now considered to be a single word. Therefore, it has been uncertain how many words include *xa.

This study searched for words containing this ***xa*** in both Seediq and Atayal through field work on Paran Seediq and investigations on dictionaries of Truku Seediq and of two Atayal dialects. Section 2.1 lists Seediq words. Section 2.2 lists Atayal words. Section 2.3 reconstructs Proto-Atayalic. It was also found that Saisiyat has some words that are cognates with Seediq or Atayal words containing *xa. Section 3 presents these words containing the bound numeral ***ha*** “one” in Saisiyat. Section 4 relates the Proto-Atayalic *xa to words meaning “ten” or “once” in other Formosan languages. Section 5 concludes that the bound numeral *xa “one” in Proto-Atayalic reflects the *əsa of Proto-Austronesian.

2. The bound numeral “one” in Atayalic languages

Section 2.1 shows Seediq data from two dialects, and Section 2.2 shows Atayal data from two dialects. Cognates in each language are reconstructed in Proto-Seediq or Proto-Atayal, respectively. Section 2.3 reconstructs the bound numeral in Proto-Atayalic.

2.1 Seediq[4]

Table 1 shows the words that are found to contain the bound numeral “one” in Seediq. The Paran Seediq data are from the author’s field notes; the Truku Seediq data are taken from Rakaw et al. (2006). For the set of words that show cognacy (the top nine words), the Proto-Seediq forms are reconstructed.

Table 1. Seediq words derived from the bound numeral *xa* “one”

Paran Seediq	Truku Seediq	Proto-Seediq
*cunka**xa***[5]	*sənəka**xa***[6]	*cənəka**xa** “the day before yesterday”
*muka**xa***	*məka**xa***	*məka**xa** “the day after tomorrow”
*ta**xa***	*ta**xa***	*ta**xa** “alone, one person”
*cu**xa**n*, *cuu**xa**n*	*səu**xa**l*	*cəu**xa**l “once in the past”
*ge**xa**n* “people of the same age”	*gə**xa**l* “relatives”	*gə**xa**l “a person of the same age”

4) Paran Seediq has the five vowels /a e i u o/ and 18 consonants /p b t d ʦ k g q s x h m n ŋ l ɾ w y/, and the diphthong /uj/ (Ochiai 2016: 19). Truku Seediq has the four vowels /a i u ə/ and three diphthongs /aw aj uj/ (Tsukida 2009: 56–62). The consonants are the same as those in Paran Seediq, excepting /ʦ/, which became /s/ in Truku Seediq. Throughout this paper, /ɾ/, /j/, and /ʦ/ are transcribed as *r*, *y* and *c*, respectively. According to Tsukida, *l* is [ɮ] and *g* is [ɣ] in Truku Seediq.

5) In Paran Seediq, the vowel after the nasal seems to be omitted.

6) This form is Pecoraro (1977: 169).

*kumu**xa**lan*	*kəmə**xa**lan*	*kəmə**xa**lan “the tens place”[7]
*ma**xa**n*	*ma**xa**l*	*ma**xa**l “ten”
*ne**xa**n*	*nə**xa**l*	*nə**xa**l “the future”
*te**xa**n* “once”	*tə**xa**l* “instantly”	*tə**xa**l “once”
*tumuka**xa*** “all day long”[8]	—	
*see**xa**n* “to have the urge to urinate”	—	
*tuxute**xa**n* “sometimes”[9]	—	
—	*mənə**xa**l* “once”	
—	*pə**xa**l* “once”	
—	*səmənə**xa**l* “only once”	
—	*su**xa**l* “the entire field”	
—	*təgə**xa**l* “to gather [relatives]”	

The following observations are made from the forms in the Table 1. First, the bound numeral is *xa* in both Paran Seediq and Truku Seediq. This is reconstructed as *xa in Proto-Seediq. Second, the *xa* is always preceded by elements that are considered to be prefixes in both Paran Seediq and Truku Seediq. In addition, these prefixes appear in various forms. Third, the *xa* in some cases appears without any suffix (the three cognate sets from the top and *tumukaxa* in Paran Seediq), whereas in all other cases, it is followed by a suffix. Furthermore, a particular suffix is used, contrary to the variation of prefixes. This suffix appears as *-n* in Paran Seediq and as *-l* in Truku Seediq.

Ochiai (2016) notes that there was a word-final *l* in early Paran Seediq, but that it has become *n* in modern Paran Seediq. The word-final *l* is retained in Truku Seediq. Therefore, for those words that have the suffix after *xa*, this suffix can be reconstructed as *-l in Proto-Seediq.

In terms of the semantics of these words derived from the bound numeral *xa*, they can be summarized by the following groups: singleness (“alone”), numerality (“ten”),[10] frequency (“once,” “once in the past,” “sometimes”), urgency (“to have the urge to urinate,” “instantly”), unity (“people of the same age,” “a relative,” and “to gather”), entirety (“all day long,” “the entire field”), expressions of the day (“the day before yesterday,” “the day after tomorrow”),[11] and futurity (“the future”). Among

7）This form is derived from the base *maxal “ten.

8）This form is likely to be derived from the base *kaxa, which is commonly seen in *mu-kaxa “the day after tomorrow” and *cənə-kaxa “the day before yesterday.”

9）This form is derived from the base *texal “once” (*texan* in modern Paran Seediq) by reduplication of the two initial consonants of the base, i.e., early Paran Seediq *təxə-təxal* > *tuxu-texan*.

10）The word for “the tens place,” *ku-muxal-an* in Paran Seediq and *kə-məxal-an* in Truku Seediq, is derived from the numeral “ten” through affixation of the circumfix *ku-…-an* or *kə-…-an*. In these forms, the first vowel *a* in the base “ten” is changed to *u* or *ə*, which represents a weakened vowel. Regarding the relation between “one” and “ten,” Ogawa (1944: 492) pointed out that the forms for “ten” derived from “one” in some Austronesian languages originally meant “to do once,” which was semantically extended to mean “having counted one cycle of numbers.” A cycle of numbers in these languages constitutes “ten.”

11）Ochiai (2020: 10) pointed out that “the day before/after yesterday” is conceptually easy to relate to *xa* “one,” if they are considered as “one day skipped from today.”

these, urgency may be difficult to semantically relate to the numeral "one." Regarding this, Takamizu (1999: 78) pointed out that "one" is related to smallness of degree or quantity and to shortness of time by providing examples from Japanese, such as *hitokoto* "a few words" (一言), *ichio kiite juo shiru* "to know the whole from a single bit of information" (一を聞いて十を知る), *isshō ni fusu* "to dismiss a matter with a laugh" (一笑に付す), *ikkasei* "temporary" (一過性), and *ikkō* "(to give) a thought" (一考).[12] The meaning of futurity could be semantically extended from the expected original meaning of "the next time, the following occasion."

In Seediq, the stress falls on the penultimate syllable, and the vowels in the pre-stress syllables are weakened. Weak vowels appear as *ɯ* in Paran Seediq (Asai 1953: 7–8, Ochiai 2016: 29, 114–115), and as *ə* in Truku Seediq (Tsukida 2009: 110–111).[13] This is the reason why all the pre-stressed vowels are reconstructed in this paper as the weak vowel *ə* in Proto-Seediq forms.[14]

2.2 Atayal[15]

Table 2 shows the words that are found to contain the bound numeral "one" in Atayal. The Squliq Atayal is from Ogawa (1931), and the C'uli' Atayal is from the Council of Indigenous Peoples (2013). For the set of words that show cognacy (the top five words), the Proto-Atayal forms are reconstructed. For the C'uli' Atayal words, their subdialect is specified in brackets because the forms in this dialect are rather diverse across subdialects in some cases.

Table 2. Atayal words derived from the bound numeral *xa* "one"[16]

Squliq Atayal	C'uli' Atayal	Proto-Atayal
məkaxa "the day after tomorrow"	*makaxa?* (Palngawan) *makaxa?* (Mayrinax)	*məka**x**a* "one day skipped from today"

12) Takamizu (1999: 79) also pointed out that the numeral "one" semantically extends to mean the whole, singleness, and identity. This is similar to the meaning of unity seen in the Seediq bound numeral "one."

13) To be precise, Asai (1953) says the weak vowel in Paran Seediq appears as /ɯ/.

14) The weak vowel in Proto-Seediq is reconstructed as *ə* instead of *ɯ* because *ə* is typical as a weak vowel, and Paran Seediq had *ə* historically.

15) Based on my field notes, Squliq Atayal has the following phonemes: the vowels /a e i o u ə/, and the consonants /p β t k ɣ q ʔ s x h ʐ r l m n ŋ y w/. The consonants β and ɣ are written as *b* and *g* orthographically. Among these vowels, *e* and *o* are observed to date back to the diphthongs *ay* [aj] and *aw*. According to Huang (1995:16–17), C'uli' Atayal has the same phoneme inventories except that it adds /ʦ/ (written as *c* in this paper) and lacks the /ə/.

16) Atayal also follows the phonological rule of vowel weakening in pre-stressed syllables (Ogawa and Asai 1935: 21), similar to the rule in Seediq. Ogawa and Asai (1935: 22) describe that the stress as falling on the penultimate syllable unless the penultimate syllable is a schwa or the word-final syllable ends in a long vowel or a glottal stop. Therefore, in the forms reconstructed as Proto-Atayal, the vowels in the pre-stress syllables are shown with the schwa. The weak vowel seems to have changed to *a* in some subdialects of C'uli' Atayal, as observed by Huang (2018: 273–274). However, in *cumukaxa* "the day before yesterday" (the Mayrinax subdialect), it is uncertain why the vowel in the initial syllable, i.e., the fourth-to last, is *u*.

*məkax**a** kəwayal* “the day before yesterday”	*makax**a**ʔ* (Palngawan) *cumakax**a**ʔ* (Mayrinax) *kax**a** pilow* (Hanxi) *səkax**a*** (Mekarang) *cəkax**a*** (Skikun)	*məkax**a*** “one day skipped from today”
*səx**a**, sax**a***[17]	*cax**a**ʔ* (Mayrinax)	*cax**a*** “alone”
*sux**a**n*	*cux**a**n* (Palngawan)	*cux**a**n* “tomorrow”
*minx**a**l, mənx**a**l* “once”	*minnəx**a**n* “again” (Skikun)[18]	*minx**a**l* “once”
*inx**a**lan məhaŋal* “one piece of luggage”[19]	—	
*kinsux**a**n* “the following day”	—	
*mənx**a**l* “last time”		
*məsinx**a**l* “to wear one more article of clothing”		
*məsinx**a**l binkagan* “to line up in a row”[20]		
*məkasx**a**l maʔabi* “sleep for a short while”[21]		
*minex**a**l* “frequently”		
*pətəgəx**a**l inənubu* “an epidemic”[22]		
*sipətəx**a**l məhuqil* “to be killed instantly”[23]		
*təx**a**l* “once”		

The following observations are made from the forms in Table 2. First, the bound numeral is *xa* in both Squliq Atayal and C’uli’ Atayal. This is reconstructed as *xa in Proto-Atayal. Second, the *xa* is always preceded by elements that are considered to be prefixes in both Squliq Atayal and C’uli’ Atayal. In addition, these prefixes appear in various forms. Third, the *xa* in some cases appears without any suffix, as seen in the top three cognate sets, while in all other cases it is followed by a suffix. Furthermore, a particular suffix is used (excepting the fourth cognate set, for which see below), contrary to the variation of prefixes. This suffix appears as *-l* in both Squliq Atayal and C’uli’ Atayal; therefore, it can be reconstructed as *-l in Proto-Atayal. It turns out that these three observations are similar to those of Seediq.

17） The second form, *saxa*, is from the Council of Indigenous Peoples (2013).

18） In this Atayal form, the final consont is *n*. However, the expected consonant is *l*. It is likely that the final *l* changed to *n* in this form. Also, in this form, the consonant *n* seems to undergo gemination (each segment belongs to different syllable). The prefix would be *min-* because the parallel word formation is seen in *min-qutux* “once,” derived from the root *qutux* “one” also in the Skikun subdialect.

19） The first word including *xa* is followed by the suffix *-l*, and is then surrounded by a circumfix *in-...-an*. This circumfix is likely to be an allomorph of < *in* >*...-an*, the past tense of the undergoer voice marker for the location subject, as is reported in Rau (1992: 36). The second word *məhaŋal* means “to shoulder” (Ogawa 1931: 84).

20） The second word *binkagan* means “a line, a row” (Ogawa 1931: 410).

21） The second word *maʔabi* means “to sleep” (Ogawa 1931: 287).

22） The second word *inənubu* means “sickness” (Ogawa 1931: 318).

23） The second word *məhuqil* means “to die” (Ogawa 1931: 164).

The segment *s* in Squliq Atayal appears as *c* in C’uli’ Atayal in the third and fourth cognate sets. For these cases, *c is reconstructed in Proto-Atayal based on the analysis of Li (1981: 260) that the Proto-Atayal *c becomes *s* in Squliq Atayal. In addition, for words ending in a vowel in the Palngawan and Mayrinax subdialects of C’uli’ Atayal as seen in the first and second examples from the top, the glottal stop is added after the vowel, which can be regarded as a phonetic realization. Therefore, it is not considered a suffix.

The fourth cognate set is peculiar in that the suffix appears to be *-n*, although the expected suffix for the word derived from *xa* is *-l*. In this case, the suffix attached to *xa* could be *-an*, which is a nominal marker for spatiotemporal expressions. The earlier form could be *su-xa-an, and the similar vowels could have coalesced, resulting in *su-xa-n. A parallel example is seen in *caman* “morning” in Paran Seediq.[24] Regarding this Seediq form, Ogawa (1939: 8–9) showed cognates in other Austronesian languages without the final consonant *n,* e.g., *ma-dama* in Siraya, and analyzed the Seediq form as being derived from the suffixation of *-an* followed by the deletion of the similar vowel, i.e., *cama-n* < *cama-an*.[25]

If this analysis is on track, Atayal and Seediq differ in the use of *xa* for days after today. In Atayal, both “tomorrow” (i.e., Proto-Atayal *suxan) and “the day after tomorrow” (i.e., Proto-Atayal *məkaxa) derive from *xa*. The two are differentiated by the use of distinct prefixes. In Seediq, only “the day after tomorrow” (i.e., Proto-Seediq *məkaxa) derives from *xa*.

For “the day before yesterday,” Atayal seems to use the same form as “the day after tomorrow” (i.e., Proto-Atayal *məkaxa). Regading C’uli’ Atayal, the Palngawan subdialect uses the same form—*makaxa*—for both meanings. The Mayrinax subdialect uses *makaxa* for “the day after tomorrow,” and a form derived from this base, *cu-makaxa*, is used for “the day before yesterday.” The addition of a prefix signifies the past. Such prefixation is seen also in the Mekarang subdialect (i.e., *sə-kaxa*) and in the Skikun (i.e., *c-kaxa*) subdialect. Squliq Atayal uses *məkaxa* for “the day after tomorrow” and a compound *məkaxa kəwayal* for its past. The second word of the compound is probably derived from *wayal*, which means “to leave.”[26] The same compounding strategy is seen in the Hanxi subdialect (C’uli’ Atayal), in which *kaxa* is followed by another word, *pilow*.[27] Some C’uli’ Atayal forms (Mekarang, Skikun, and Hanxi) omit the word-initial *mə-.

In summary, it seems that *məkaxa was used indifferentially for “the day before yesterday” and “the day after tomorrow.” It might mean “one day skipped from today, whether past or future.” Later, this form tended to be used for “the day after tomorrow,” and a prefix or another word is introduced when specifying “the day before yesterday,” except in the Palngawan subdialect (C’uli’ Atayal). Because the additional element was a later innovation in each dialect, its adoption was not uniform. Some subdialects use prefixation (Mayrinax, Mekarang, Skikun subdialects of C’uli’ Atayal), whereas others use compounding (Squliq Atayal and the Hanxi subdialect of C’uli’ Atayal).

24）However, *caman* is archaic form in modern Paran Seediq. The meaning given for it by the author’s informants is “dawn.”

25）This discussion is detained in Ochiai (2019).

26）This form is taken from Ogawa (1931: 24).

27）The exact meaning of this word is uncertain. However, it is inferred to have the meaning of past-ness.

Furthermore, the prefixes or the additional word in the compound are not uniform among dialects.[28] This shows that these developed later in each (sub)dialect.

These Atayal words that derive from the bound numeral *xa* are similar to those in Seediq in terms of semantics too. These Atayal words can be semantically categorized into the following groups: singleness (“alone,” “one piece,” “one more piece,” “one line”), frequency (“once,” “frequently”), urgency (“instantly,” “short while”), time expressions (“last time”), expressions of the day (“tomorrow,” “the day after tomorrow,” “the following day,” “the day before yesterday”).

It is uncertain what *pətəgəxal* means exactly in the compound *pətəgəxal inənubu* “epidemic,” whose second word means “sickness.” However, *pətəgəxal* could be a cognate with the Truku Seediq *təgəxal* “to gather.” The Atayal form seems to have the additional prefix *pə-*, which is likely to be a causative prefix. This compound probably means “causing to be entirely sick (everyone gets sick).”

2.3 Proto-Atayalic reconstruction

In Section 2.1, the bound numeral “one” is reconstructed as *xa in Proto-Seediq based on *xa* in Paran Seediq and *xa* in Truku Seediq.

For Squliq Atayal, Ogawa (1944: 469–470) says that the *ha* in *min-ha-l* “once” (Squliq Atayal) is a morpheme meaning “one.” However, all the examples of this bound numeral “one” extracted from his Squliq Atayal dictionary (Ogawa 1931) have *xa* instead of *ha*, as seen in Table 2. The form of the bound numeral “one” is the same in C’uli’ Atayal. Therefore, *xa is reconstructed in Proto-Atayal in Section 2.2.

Then, the bound numeral “one” in Proto-Atayalic can be reconstructed as *xa on the basis of the Proto-Seediq *xa and the Proto-Atayal *xa.

Among the words that derived from *xa, *məkaxa would safely be reconstructed in Proto-Atayalic because this is reconstructed to both Proto-Seediq and Proto-Atayal. Regarding its meaning, it might mean “one day skipped from today, whether past or future” as discussed in Section 2.2.

Another form that can be reconstructed in Proto-Atayalic is *təxal “once.” Although the reflex of this form was not obtained for C’uli’ Atayal, the identical form, *təxal*, is seen in Squliq Atayal as the last example in Table 2. In addition, as in the ninth example in Table 1, the identical form, *təxal, is reconstructed in Proto-Seediq.[29]

When there is a suffix following *xa*, only one particular suffix is used in Seediq and in Atayal. This suffix is reconstructed as *-l in Proto-Seediq and as *-l in Proto-Atayal. Based on the above, the Proto-Atayalic form is reconstructed as *-l. According to Zeitoun et al. (2010: 875–876), the suffix *-N is

28） As for prefixes, *cu-* in Mayrinax, *cə-* in Skikun, and *sə-* in Mekarang are likely to originate from the same form. The Proto-Seediq *cənə-, which could be further segmented as *c<ən>ə-* with the infix indicating past (seen in *cənəkaxa), may be the related prefix. This prefix could be reconstructed as the Proto-Atayalic *cə-. However, the bases used for the derivations are different. They are *makaxa* in Mayrinax, but *kaxa* in Skikun, Mekarang, and Proto-Seediq.

29） The Proto-Atayal *caxa and the Proto-Seediq *taxa both meaning “alone, one person,” are likely to be cognates. However, the word-initial consonants are difficult to reconstruct because the *c in Atayal and the *t in Seediq show an irregular correspondence.

reconstructed in Proto-Austronesian with the meaning of "recurrence," which indicates a number of times, such as "once," "twice," and "three times." They provide data for *-N from several Formosan languages including Paiwan (e.g., *paka-pusa-**l*** "for two days"), Tsou (e.g., *i?-po-pso-**hɯ*** "twice"), Rukai (e.g., *?o-a-**le*** "once"), Puyuma (e.g., *paria-sa-**l*** "once"), Saisiyat (e.g., *mon-ha-**l*** "once"), Thao (e.g., *k<un> turu-**z*** "to eat three times"), and Bunun (e.g., *ma-pit-s?a-**n*** "to cook once").[30] Their data do not include Atayalic languages. This paper supplements the Proto-Atayalic *-l as the reflex of the Proto-Austronesian *-N.

3. The bound numeral "one" in Saisiyat[31]

A handful of words that derive from the bound numeral "one" in similar fashion to Atayalic languages are found in Saisiyat, among other Formosan languages. This section considers Saisiyat in the view of its bound numeral "one" and its relation to Atayalic languages.

Zeitoun et al. (2015: 162) explains that the bound numeral "one" is *ha* in Saisiyat. This bound numeral appears as *hae* [hæ] when followed by a glottal stop.[32] The bound form must be a shortened form of the numeral *?aehae?* "one."[33] Zeitoun et al. (2015) provide three examples of the bound numeral *ha*, including *(h)inhal* "to beat once," *monhal* "once," and *kamonhal* "last year."[34]

Similar to Atayalic languages, it is possible that for speakers of the Saisiyat of the present day, the origin of *ha* is forgotten. It is uncertain how many words include *ha*. Therefore, the present study searched the data of Zeitoun et al. (2015) for words that may include this covert *ha*. Table 3 shows an exhaustive list, including the three examples from the previous paragraph. One example is added from Ogawa (1939: 15) and another is added from Li (1978), as indicated in the footnotes. Any cognates found in Proto-Seediq or Proto-Atayalic are shown in the middle and the right-hand columns, respectively.

Table 3. Saisiyat words derived from the bound numeral *ha* "one"

Saisiyat	Proto-Seediq	Proto-Atayal
*ka**ha**-kare?al*[35] "the day before yesterday"	*məkaxa	*məkaxa
*?apnae**hae?*** "each one"		

30) According to Blust (2013: 582), the Proto-Austronesian phoneme *N is reflected in various phonemes in the daughter languages: *l*^{y} in Paiwan, *h* in Tsou, *z* in Thao, *n* in Bunun, and *l* in Puyuma, Saisiyat, Atayal and Seediq. In the forms given as examples of *-N, the Paiwan form should be understood as having the final consonant *l*^{y}. The Tsou and Rukai forms have epentetic vowels following the final consonant, *h* and *l* respectively, in order to phonotactically avoid a word ending in a consonant.

31) According to Zeitoun et al. (2015: 15–16), Saisiyat has the following phnonemes: the consonants /p β t k ʔ θ ð ʃ h m n ŋ r l w j/ and the vowels /i ə æ œ o a/. In this paper, /θ/, /ʃ/, β/, /ð/, /j/, /ə/, /æ/, and /œ/ are represented as *s*, *sh*, *b*, *z*, *y*, *e*, *ae*, and *oe*, respectively.

32) In the Saisiyat data from Zeitoun et al. (2015), words ending in a vowel are, with a few exceptions, followed by a glottal stop.

33) This form is taken from Zeitoun et al. (2015: 161).

34) "Last year" can be understood as distancing "one" year from the present year.

35) This form is taken from Ogawa (1939: 15).

*tae?**hae**?*, *hae?**hae**?* “nine”[36]		
*soe**hae**?* “to take one”		
*yoe?**hae**?* “seven”[37]		
*?ae**ha**el* “a sibling/cousin”	*gəxal	
*?aeh?**ha**el* “a relative”		
*(h)in**ha**l* “to beat once”		
*kamon**ha**l* “last year”[38]		
*koe?**ha**l* “next year”		
*ma?oen**ha**l* “to go together”		
*may**ha**l* “later”		
*mina:**ha**el* “offspring”[39]		
*min**ha**l* “to bloom once a year”		*minxal “once”
*mon**ha**l* “once”		
*main**ha**l* “to work for each other”		
*tatilhae**ha**l*[40] “to help”		
*tatoe?ae**ha**l* “to treat as the same generation”		
*tinoe?ae**ha**l* “a relative of the same generation”		
*shanhae**ha**l* “to help eat”		
*nae**han***[41] “again”	*nəxal “future”	

Regarding the first word in Table 3, Ogawa (1939: Appendix) recorded the word *kare?al* meaning “the day after tomorrow,” which seems identical to the latter half of *kaha-kare?al* “the day before yesterday.” The word-initial segments *kaha* could correspond to *məkaxa “one day skipped from today, whether past or future” in Atayalic languages, as *kaxa* is used as a base for other derivations in some cases. The corresponding form, *kaha*, is perhaps retained in Saisiyat as the compound *kaha kare?al*.

36）The reason that “nine” contains “one” is that “ten” is the base for “nine,” i.e., 10 minus 1, in Saisiyat as pointed out by Li (2011: 201).

37）The reason that “seven” contains “one” is that “six” is the base for “seven” in Saisiyat. As Ogawa (1944: 486, appendix) points out, the form for “seven” derives from “six” in Saisiyat: “six” is *saiβusiR* and “seven” is *saiβusiR-o-aha*, which signifies 6 plus 1. Regarding “seven,” Zeitoun et al. (2015: 159) note that *shayboshi (:)=o ?aehae?*, which corresponds to Ogawa’s form for “seven,” is used as a cardinal number, whereas *yoe?hae?* is used in serial counting. Both *yoe?hae?* and *shayboshi(:)=o ?aehae?* have the word-final *hae* “one.” It is likely that the former is derived from the latter; however, the origin of the initial syllable *yoe?* in *yoe?hae?* is unknown.

38）It is likely that *monhal* “once” is the base for this form.

39）This form is taken from Li (1978: 158). The vowel after *h* is written as *æ* in this source, whereas the vowel under this condition is written as *a* in Zeitoun et al. (2015).

40）The segment *hae* preceding the *ha* “one” could be a reduplication of *ha*, which probably indicates reciprocity.

41）In this Saisiyat form, the final consont is *n*. The expected consonant is *l*, just as it is in the cognate in Proto-Seediq. It is likely that the final *l* changed to *n* in this form. However, this change is only seen in this example among the words in Table 3.

Similar to Atayalic languages, the bound numeral *ha* “one” in Saisiyat must accompany prefixes, and these prefixes appear in various forms. The *ha* either appears without a suffix, as seen in the first element of a compound in the first example, or it appears with a word-final glottal stop, as seen in the second to fifth examples. In all other cases excepting the last example, *ha* is followed by the particular suffix *-l*, which is a reflex of the Proto-Austronesian *-N as discussed in Section 2.3.

In addition, these Saisiyat words are semantically similar to those of Atayalic languages. The meanings of these words can be categorized into the following groups: singleness (“each one,” “to take one,” “to beat once,” “to bloom once a year”), numerality (“seven,” “nine”), unity/collectiveness (“a relative,” “offspring,” “a relative of the same generation,” “to go together,” “to treat as the same generation,” “to work for each other,” “to help eat”), expressions of the day (“the day before yesterday”), expressions of adjacent years (“last year,” “next year”) and futurity (“later,” “again”).

For *ʔaehael* and *naehan* in Saisiyat, *gəxal and *nəxal are provided in Proto-Seediq as their cognates.[42] However, there is a problem regarding the correspndence between the penultimate vowel. The *ə in Proto-Seediq reflects the *ə in Proto-Austronesian, which in turn is reflected as the *ə* in Saisiyat.[43] Therefore, the expected forms for Saisiyat are **ʔəhael and **nəhal, respectively.[44]

It is possible that the *ə* in Saisiyat changed to *ae* when followed by *h*. This change is seen in *ʔaehae?* [ʔæhæʔ] “one,” which reflects the Proto-Austronesian *əsa, and the vowel *ə in Proto-Austronesian is expected to appear as *ə* in Saisiyat. However, it appears as *ae*. This vowel is followed by the consonant *h*.

Another example is suggestive of this change. The Saisiyat word *sibaeLaeh* [ʃiβæʎæh][45] “to borrow” has a cognate in Paran Seediq: *k<um>barux* “to borrow/lend labor.” Their roots *baeLaeh* in Saisiyat and *barux* in Seediq, are cognates. The Seediq form is inflected as *kuburex-i* (imperative). In this inflected from, the vowel in the final syllable *u* becomes *e* when it moves to the penultimate syllable by the addition of a suffix. The vowel that exhibits this phonological change originates in the historical *ə*, according to Ochiai (2018). In Paran Seediq, the historical *ə* was fortified and became *e* in the penultimate syllable. Therefore, the Proto-Seediq form is *barəx. It follows that in the Saisiyat form, the vowel in the final syllable is expected to be *ə*; however, it appears as *ae*. This could also be an *ə* in Saisiyat that changed to *ae* due to the influence of the following consonant *h*.

In form and meaning, the words derived from the bound numeral *ha* “one” in Saisiyat show a strong connection with those of Atayalic languages. It is almost certain that the *ha* in Saisiyat is a cognate of the *xa in Proto-Atayalic. The remaining issue is how the proto-form of these two is reconstructed. According to Ross (2015: 31–32), the phoneme *h* in Saisiyat dates back to four

42) Proto-Seediq *g goes back to the Proto-Austronesian *R, which in turn is reflected as /ʎ/ in the Taai dialect and as zero in the Tungho dialect of Saisiyat, according to data from Li (1978). In the case of *ʔaehael* “sibling/cousin,” which is from data of Zeitoun et al. (2015) on Tungho Saisiyat, the reflex of *R appears as zero but a vowel is preceded by a glottal stop.

43) Li (1978: 140) says that the Proto-Austronesian *ə is reflected as *ə* in Saisiyat unless it is contiguous to *q or *s.

44) The double asterisk indicates that these forms are not observed.

45) This form is from Li (1978: 182).

phonemes in Proto-Austronesian: *s, *θ, *x, and *h. However, in Ross's table of sound correspondence, none of these proto-phonemes are reflected as *x* in Proto-Atayalic. Moreover, there was no phoneme *x* in Proto-Atayalic. Regarding the Atayalic *x*, Li (1981: 263) notes that its origin is unknown.

To resolve this issue, the next section searches for cognates of the Proto-Atayalic *xa in other Formosan languages. Some Formosan languages have words that show cognacy with the Proto-Seediq *maxal "ten."

4. The proto-form *ma-sa-N "ten"

Ogawa (1944: 492) explains that the numeral "ten" in some Formosan languages derives from *əsa, which is identical to the numeral "one" in Proto-Austronesian, and he gives this proto-form as *ma-sa-əN.[46] This proto-form shows that the root *əsa "one" is shortened to *sa.

Ogawa (1944: 469–470) points out that this *sa is retained as *xa* in the numeral "ten" in Seediq. He provids examples of the numeral "ten" of the same word formation from some Formosan languages. These examples are shown in Table 4. Based on the cognate set in Table 4, the Proto-Austronesian is reconstructed as *ma-sa-N. This is almost the same form as Ogawa's reconstruction, *ma-sa-əN, but the schwa is eliminated.

Table 4. "Ten and "twenty" derived from bound numerals in some Formosan languages[47]

	10	20
Proto-Seediq	*ma-xa-l	*mə-pusa-l
Bunun[48]	*ma-sʔa-n*	*ma-pusa-n*
Tsou	*ma-s-kə*	*ma-pus-kə*

46) The actual form Ogawa (1944) reconstructed is *ma-sa-ən₂*, whose final consonant *n₂* corresponds to *N.

47) The data in this table are from Ogawa (1944: 490, appendix), excepting Proto-Seediq, which is reconstructed in this paper.

48) The Proto-Austronesian phoneme *s is reflected as *c in Proto-Bunun (Ross 2015: 32), which is still retained as *c* in the Northern dialect of Bunun, whereas it changed to *s* in other dialects. In the Bunun form *ma-sʔa-n* that Ogawa recorded, the *s is reflected as *s*; therefore, this form belongs to dialects other than the Northern dialect. The corresponding form in Takibakha village, which belongs to one of the sub-dialects of the Northern dialect, is *macʔan* "ten" [ˈmaʦ.ʔan], according to my field notes (the informant was Mr. Atul Manqoqo), which is likely to be historically derived from ma-*cʔa-an* by a shortening of the sequence of like vowels. A word for "siblings, brethren" in this sub-dialect is *tecʔaan* [teʦ.ˈʔa.an/teʦ.ˈʔa:n], which is historically derived from *ta<i>-cʔa-an*, in which the *ta* is a kind of prefix, the < *i* > is an infix indicating the past tense, and the diphthong *ai* became *e*. (This sub-dialect has the accent on the penultimate syllable; however, the like vowels in different syllables in this example phonetically become one syllable composed of a long vowel.) This *tecʔaan* is also derived from the bound numeral "one," which appears to be *cʔa* in Proto-Bunun, which reflects the Proto-Austronesian *sa; however, the glottal stop after *c is unexpected. It is also uncertain why the sequence of like vowels was shortened to one in *macʔan* but was not shortened in *tecʔaan*. The relationship between the word for "siblings, brethren" in Bunun and the bound numeral "one" discussed in this paper were also pointed out by an anonymous reviewer.

Kanakanavu	*ma-a-nə*	*ma-pusa-nə*
Saaroa	*(k-)m-a-* l̥*ə*	*ma-puwa-*l̥*ə*
Rukai	*məŋ-əa-l*	*ma-ʔosa-lə*
Thao	—	*ma-pusha-z*
Puyuma	—	*maka-butaʔ-an*
Proto-Austronesian	*ma-sa-N	*ma-puSa-N[49]

Examples of the numeral "twenty" in these Formosan languages are shown in the right-hand column for ease of comparison. For "twenty," similar examples from two other Formosan language (Thao and Puyuma) are added.

Zeitoun et al. (2010) reconstructed the Proto-Austronesian word for "twenty" as *ma-puSa-N, and said that puSa "two" is a bound numeral. The suffix appears in various segments across Formosan languages; however, they say these are the reflex of the proto-suffix *-N "recurrence."

The words for "ten" are likely to have been derived in the same fashion as "twenty." In most cases, the prefix used in these derivations is *ma-*, and the suffix is the identical one used for "twenty" in each language. Then, the proto-form can be reconstructed as *ma-sa-N based on these cognates.

Nonetheless, the author dares not say these are Proto-Austronesian forms, because as suggested by Ochiai (2015), the Proto-Austronesian numeral system may have had four numerals only: one, two, three and four. As Ochiai (2020) says "five" in most Austronesian languages emerged by the semantic shift from "hand" and many higher numerals from "six" to "nine" in Formosan languages shows forms derived from the lower numerals from "one" to "four." Table 4 shows that the forms for "ten" in some Formosan languages are dependent on "one." The decimal system is incomplete in Formosan languages, excepting the numerals from "one" to "four."

This paper maintains that the "ten" and "twenty" derived from the bound forms of "one" and "two" are later innovations in these languages. It is unclear whether this innovation happened in a particular subgroup larger than Proto-Atayalic but smaller than Proto-Austronesian which is not known to date, whether it happened independently in each language, or whether the innovation in a particular language is borrowed by another language at a sufficiently early time when the sound changes from Proto-Austronesian were still small.

5. Conclusion

Section 4 showed that the Proto-Atayalic *xa reflects Proto-Austronesian *sa. This is in accordance with *ha* in Saisiyat (Section 3), because the Proto-Austronesian *s is reflected as *h* in Saisiyat. Still, the problem of **x* in Proto-Atayalic remains unsolved. According to Ross (2015), the Proto-Austronesian *s is reflected not as *x* but as *h* in Proto-Atayalic.

It is possible that in Proto-Atayal, the proto-phoneme *s changed in steps, first becoming *x* and then became *h*. This phoneme appears as *h* in modern Atayalic languages in most cases (e.g., Proto-Austronesian *saləŋ "pine tree" > Proto-Atayalic *haruŋ).[50] However, it might remain as *x*, its medial

49) This form is reconstructed in Proto-Austronesian in Zeitoun et al. (2010).

stage of the sound change, in the case of the bound numeral *sa.[51)]

In conclusion, the Proto-Atayalic *xa is a bound numeral meaning "one" that reflects the shortened form of the Proto-Austronesian *əsa "one." The Proto-Atayalic *xa is a bound numeral that requires derivation by affixation. It has to accompany a prefix, which appears in various forms. The *xa appears with or without a suffix. The suffix directly following *xa is *-l, which reflects the Proto-Austronesian *-N "recurrence." Atayalic words derived from the bound numeral share a few cognates with Saisiyat, in which the bound numeral appears as *ha*.

References

Asai, Erin (1953) *The Sedik Language of Formosa*. Kanazawa: Cercle Linguistique de Kanazawa.

Blust, Robert (1999) Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In: Elizabeth Zeitoun and Paul Jen-kuei Li (eds.) *Selected papers from the eighth international conference on Austronesian linguistics*, 31–94. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.

Blust, Robert (2013) *The Austronesian languages*. Canberra: Australian National University, Research School of Pacific and Asian Studies.

Blust, Robert and Stephen Trussel. (2010) Austronesian comparative dictionary, Web Edition. http://www.trussel2.com/ACD/, accessed May, 2020.

Council of Indigenous Peoples (2013) *Yuanzumin zuyu e leyuan*. http://web.klokah.tw/ [accessed May 2020].

Li, Paul Jen-kuei (1978) A comparative vocabulary of Saisiyat dialects. *Bulletin of the Institute of History and Philology, Academia Sinica* 49: 133–199.

Li, Paul Jen-kuei (1981) Reconstruction of Proto-Atayalic phonology. *Bulletin of the Institute of History and Philology, Academia Sinica* 52(2): 235–301.

Li, Paul Jen-kuei (2011) *Taiwan nandao Minzu de zuqun yu qianxi* [The groupings and migrations of Taiwan's Aboriginal peoples]. Taipei: Avanguard Publishing.

Huang, Lillian M. (1995) *A study of Mayrinax syntax*. Taipei: Crane.

Huang, Hui-chuan J. (2018) The Nature of pretonic weak vowels in Squliq Atayal. *Oceanic Linguistics* 57(2): 265–288.

Ochiai, Izumi (2015) Origin of the numeral "five" in Northwestern Formosan languages: Revision of

50) The Proto-Austronesian form is from Blust and Trussel (2010), and the Proto-Atayalic is from Li (1981: 290).

51) Incidentally, it was not so difficult to spot words containing *xa* in Seediq and Atayal. In relation to this, Ogawa (1931: 6–7) noted that *h* and *x* in Atayal are different phonemes but that *x* is seen in only a small number of words such as *suxan* "tomorrow," *səxa* "alone," and *minxal* "once." According to my observations, the segment *x* in both Atayal and Seediq was rarely seen in a word-medial position and almost never in a word-initial position. If there is an *x* in a word-medial position, it stands out. Almost every segment *x* in a word-medial position that the author found in the dictionary belonged to the bound numeral *xa. The origin of the word-final *x* in Atayalic languages needs to be investigated in a future study.

the Proto-Austronesian numeral system. Presented at the 25th Annual Meeting of the Southeast Asian Lingustic Society, Payap University, Chiang Mai, May 27–29.

Ochiai, Izumi (2016) Sedekkugo Paran hōgen no bunpō kijutsu to hiishisei settouji no hikaku gengogakuteki kenkyū [A descriptive grammar of Paran Seediq and a comparative study of non-volitional prefixes], Ph.D diss, Kyoto University.

Ochiai, Izumi (2018) Historical reduplication in Seediq. Kyoto University Linguistic Research 37: 23–40.

Ochiai, Izumi (2019) "Morrow" in Seediq. *The Kobe Gaidai ronso* 70(1): 131–144.

Ochiai, Izumi (2020) Numerals and related expressions in Paran Seediq, *Journal of Language and Culture of Hokkaido* 18: 1–18.

Ochiai, Izumi (2020) Proto-Austronesian *lima revisited: From archaic "hand" in Atayalic languages. *Language and Linguistics in Oceania* 11: 1–18.

Ogawa, Naoyoshi (1931) *Atayaru goshū* [Atayal word lists]. Taihoku: Taiwansōtokufu.

Ogawa, Naoyoshi (1939) *Toki ni kansuru takasagozoku no go* [Various phrases of time in the native languages of Formosa]. Japanese Journal of Ethnonogy 5(1): 1–14.

Ogawa, Naoyoshi (1944) Indoneshiago ni okeru taiwan takasagogo no ichi [The place of the Formosan languages in the Indonesian languages]. In *Taiheiyô* vol. 1, edited by Hirano Yoshitaro, 451–502. Tokyo: Kawade Shobo.

Ogawa, Naoyoshi and Erin Asai (1935) Gengo ni yoru taiwan takasagozoku densetsushū [*The Myths and Traditions of the Formosan Native Tribes*]. Taipei: Taihoku Imperial University.

Pecoraro, Ferdinando (1977) *Essai de dictionnaire Taroko-Français*. Cahier d'Archipel 7. Paris: Société pour l'Etude et la Connaissance du Monde Insulindien.

Rakaw, Lowsi, Jiru Haruq, Yudaw Dangaw, Yuki Lowsing, Tudaw Pisaw, and Iyuq Ciyang ed. (2006) *Tailugezuyu jianyi zidian* [Concise Dictionary of Taroko Seediq]. Xiulin: Xiulin Township Administrative Office.

Rau, Der-Hwa Victoria (1992) A grammar of Atayal. Doctoral dissertation. Cornell University.

Ross, Malcolm (2015) Some Proto Austronesian coronals reexamined. In: Zeitoun, Elizabeth, Stacy Fang-Ching Teng, and Joy J. Wu (eds.) *New advances in Formosan linguistics*, 1–38. Asia-Pacific Linguistics, 2015.

Takamizu, Toru (1999) Susini mirareru imikakucho keikenteki kibanga gengoni hanei sarerukotono ichirei [Semantic extensions of numerals: Expressions through experiential basis]. *Papers in Linguistic Science* 5: 75–87.

Tsukida, Naomi (2009) Sedekkugo (Taiwan) no bunpo [A grammar of Seediq (Taiwan)]. Doctoral dissertation, The University of Tokyo.

Zeitoun, Elizabeth, Stacy, Fang-ching Teng, and Raleigh Ferrell (2010) Reconstruction of '2' in PAN and related issues. *Language and Linguistics* 11(4): 853–884.

Zeitoun, Elizabeth, Tai-hwa Chu, and Lalo a takesh kaybaybaw (2015) A study of Saisiyat morphology. Honolulu: University of Hawaii.