

Formosan “leaf”: A reconstruction

Izumi OCHIAI

1. Introduction

This study reexamines and reconstructs the words for “leaf” in Proto-Austronesian (PAN) by investigating Formosan languages. Formosan languages are spoken by the indigenous population in Taiwan, except for Yami, which belongs to the Malayo-Polynesian languages. More than 20 Formosan languages are known, including the Atayalic subgroup (Atayal and Seediq), Tsouic subgroup (Tsou, Kanakanavu, and Saaroa), Bunun, Rukai, Paiwan, Puyuma, Amis, Kavalan, Basay (extinct), Ketangalan (extinct), Taokas (extinct), Saisiyat, Babuza/Favorlang (extinct), Pazih, Thao, Papora (extinct), Hoanya (extinct), and Siraya (extinct).

Formosan languages are crucial to the reconstruction of PAN forms because of their genealogical position in the Austronesian language family. According to Blust’s (1999a) hypothesis, Proto-Austronesian diverged into 10 subgroups. Among these first-order subgroups, all, but one, belong to the Formosan languages: the Atayalic subgroup, Tsouic subgroup, and others. The only non-Formosan subgroup is Proto-Malayo-Polynesian. Therefore, the evidence from the Formosan languages provides strong support for Proto-Austronesian reconstruction.

For example, Blust and Trussel (2010) reconstructed PAN “leaf” as *waSaw, based on the forms attested in two Formosan languages, Seediq and Paiwan. Table 1 shows Blust and Trussel’s data with minor modifications by the author. The Seediq form *wasau*, cited from Ferrell (1969), was changed to *wasaw*, according to the Truku Seediq dictionary (Rakaw 2006).

Table 1: PAN “leaf” in Blust and Trussel (2010)

PAN	*waSaw
Seediq	wasaw
Paiwan	asaw

The reconstruction is based solely on these two Formosan languages. Although the match may seem coincidental at first, the sound correspondence is impeccable. The *S becomes *s* in Seediq and Paiwan, *w remains as it is in Seediq, and *w is either retained or becomes zero in Paiwan.¹ Therefore, it is difficult to doubt this reconstruction. Regarding the Paiwan data, Ogawa and Asai (1935: Appendix 29) reported “leaf” in the Tjavuali dialect of Paiwan as *wasau*, indicating that Proto-Paiwan can be reconstructed as *wasaw. In addition, they reported *vasau* as a form in one of the Rukai dialects, collected in Tamalakau village. This word could be a loanword from Paiwan because, according to Ogawa and Asai’s (1935) map, the village is located near Paiwan villages, and other Rukai dialects do not use the form.

The aim of this paper is to propose another reconstruction for the Proto-Austronesian word for “leaf.” The form to be reconstructed is *RabaR; it has cognates in several Formosan languages, discussed in Section 2. In relation to “leaf,” Section 3 overviews the Formosan reflexes of PAN *biRaq, which is reconstructed with the meaning “Alocasia.”

2. Proto-Austronesian *RabaR

This section begins with a discussion of “leaf” and its relation to “summer” in the Atayalic subgroup. Next, the Proto-Atayalic “leaf” is reconstructed, with which the Proto-Austronesian form is hypothesized as a first step to the search for cognates in other Formosan languages.

2.1 Atayalic

Li (1981) collected the Atayalic forms for “leaf” in five Atayal villages (Squiliq, Maspazi?, Skikun, Mayrinax, and Plngawan) and four Seediq villages (Tongan, Toda,

¹ For Seediq sound correspondence, see Li (1981) (He has not reconstructed Proto-Seediq but directly reconstructed Proto-Atayalic from Seediq dialects and Atayal dialects). For Paiwan sound correspondence, see Ho (1978).

Truwan, and Inago). According to Ogawa and Asai’s (1935) classification, Atayal has two dialects, Squliq and C’uli’. Of the villages Li investigated, only Squliq belongs to the Squliq dialect, and the others belong to the C’uli’ dialect. According also to Ogawa and Asai (1935), Seediq has two dialects, Paran and Truku. Among these villages, only Tongan belongs to Paran, and the others belong to Truku.

Li listed the forms for “leaf” collected in each village, except for the cases of Tongan and Inago (Table 2). Striking similarities were observed within the Atayalic and Seediq forms, respectively, and across the two languages. However, Li did not reconstruct Proto-Atayalic based on these data (He reconstructed neither Proto-Atayal nor Proto-Seediq.) However, because of the clear sound correspondences, some hypotheses should be proposed.

Table 2: Atayalic “leaf” in Li (1981:287)

Language	Village	“Leaf”
Atayal	Squliq	<i>ʔabaw</i>
	Maspaziʔ	<i>ʔabaw</i>
	Skikun	<i>ʔabax</i>
	Mayrinax	<i>ʔabag</i>
	Plngawan	<i>ʔabaw</i>
Seediq	Tongan	---
	Toda	<i>nabaw</i>
	Truwan	<i>nabaw</i>
	Inago	---
Proto-Atayalic		---

Across Li’s (1981) Atayalic data, word-initial glottal stops demonstrate a restricted distribution. For example, they appear in the word-initial position preceding a vowel or in the word-final position following a vowel. This finding suggests that, here, the glottal stop is phonetic rather than phonemic. Then, the more precise form should be *abax* (Skikun), *abag* (Mayrinax), or *abaw* (other Atayal). The variation in the word-final position reflects the earlier word-final *g*, based on Li’s (1981:257) that the Proto-Atayal word-final *g* is retained in Mayrinax, devoiced and fricated in Skikun, and changed to *w* in other dialects including Seediq dialects. Then, Proto-Atayal is **abag*. For Seediq, we can infer that only the Proto-Seediq could be

*nabag with an extra *n* in the word-initial position. These reconstructed forms at least have the sequence *abag* in common.

Additionally, Pecoraro (1977) demonstrated that Truku Seediq (its subdialect is Inago in Li's [1981] data) has *rə nabaw* as a variant of *nabaw*. Then, in addition, the Proto-Seediq could have an extra segment *r* in front of it. This segment *r* is critical in relating “leaf” to “summer,” as discussed next.

Table 3 shows Li's (1981) data for “summer.” He reconstructed Proto-Atayalic “summer” as *-bagan, indicating in his analysis that some segments may precede it; however, these cannot be reconstructed easily. This reconstruction must be refined by the use of the conceptual connection with “leaf.”

Table 3: Atayalic “summer” in Li (1981:294)

Language	Village	“Summer”
Atayal	Squliq	<i>bagan</i>
	Maspazi?	<i>?abagan</i>
	Skikun	<i>gbagan</i>
	Mayrinax	<i>gabagan</i>
	PIngawan	<i>?abagan</i>
Seediq	Tongan	<i>rbagan</i>
	Toda	<i>rbawan</i>
	Truwan	<i>rbagan</i>
	Inago	<i>rbagan</i>
Proto-Atayalic		<i>*-bagan</i>

Utsushikawa (1936:14) asserted that the Atayal word for “summer” comprises “leaf” and “thrive.” His observation suggests that “summer” is explained as the season in which leaves grow. However, “summer” is not the composite of the Atayal words *abao* “leaf” and *bangan* “thrive,” as he said. This paper proposes that “leaf” (Proto-Atayalic *Rabag; see below) bears the suffix *-an*, through which abstract nouns are often derived from in the Austronesian language. Along this line, the final segment *-an* for “summer” (Table 3) is originally a suffix.

For example, the other season, winter, is composed by the same derivation. A hypothetical Proto-Atayalic word *qamis “cold wind” (this form is no longer used independently in Atayalic languages) is attached with *-an* (cf. PAN *qamiS “north

wind”), resulting in *qamis-an*, which is the word for “winter” in Atayal. Seediq cognate is *misan* which shows deletion of the initial syllable from *qamisan*. Atayal and underwent initial vowel deletion to become *misan* “winter” in Seediq.

Based on the Seediq forms (Table 3), the Proto-Seediq is easily obtained as *r(ə)bagan (only Toda underwent *g* to *w*). To further analyze the form diachronically, it is likely to be composed of Proto-Seediq *rabag “leaf” and the suffix *-an*. Proto-Seediq “leaf” is reconstructed with the first vowel as *a*, because it is evidently *a* in “leaf” (Table 2). Because the accent is on the penultimate syllable, the pre-stress vowel *a* underwent reduction to a schwa. In Truku Seediq, for unknown reasons, *rabag “leaf” underwent infixation of *n* after the initial consonant, resulting in *r<n>abaw* /rənabaw/ or *<n>abaw* with further deletion of the initial consonant.

In the Atayal forms, the first segment is either lost (i.e., in Squliq, Maspazi?, and PIngawan) or shown as *g*. Proto-Atayal can be reconstructed as *g(ə)bagan. Thus, “summer” is composed of Proto-Atayal *gabag “leaf” and the suffix *-an*. However, the actual form for “leaf” has no initial consonant in each Atayal dialect in Table 2. The initial *g* in “leaf” seems to be deleted sporadically in Proto-Atayal. This probably occurs because having two *g*’s is avoided by deleting one of them.

According to the present reconstruction for “leaf,” Proto-Seediq is *rabag and Proto-Atayal is *gabag. They differ in the initial consonants: *r* and *g*. This divergence presumably originates in PAN *R /ɣ/ because only these segments have the reflex *r* and *g* in Atayalic languages. Regarding this divergence, Li (1981:274) described the condition as PAN *R became *r* in front of *i* and became *g* elsewhere. However, Proto-Seediq violates this condition because *R became *r* in front of the vowel *a* (e.g., Proto-Seediq *rabag). In relation to this point, Li (1981:255) mentioned that PAN *R changed to *r* in Seediq between two *a*’s in examples such as Seediq *dara* “blood” (Li’s Proto-Atayalic is *daga?). Incorporating the data for “leaf,” this condition may be loosened to assert that in front of *a*, *R became *r* in Seediq instead of between two *a*’s. In summary, the initial consonant can be reconstructed as *R. Proto-Atayalic “leaf” then becomes *Rabag.

Next, the PAN form is hypothesized based on the Proto-Atayalic form. During the reconstruction process for Proto-Atayalic, the segment *R has already returned to the stage of PAN. Other segments, *a* and *b*, also date back to PAN. The remaining segment is *g*, which is nonexistent in PAN. The only source for this segment is, as discussed in the previous paragraph, PAN *R. Consequently, the PAN *RabaR is obtained (Table 4).

Table 4: Atayalic “leaf” and “summer”: Reconstruction

	Leaf	Summer
Proto-Atayal	*gabag	*gabag-an > <i>gəbagan</i>
Proto-Seediq	*rabag	*rabag-an > <i>rəbagan</i>
Proto-Atayalic	*Rabag	*Rabag-an
PAN	*RabaR	

The reconstruction for PAN *RabaR “leaf” is based solely on the Atayalic subgroup. To establish its status, more cognates from other Austronesian languages, from Formosan languages, are necessary.

2.2 Pazih

A cognate in Pazih *rabax* “leaf” was easily observed (Li and Tsuchida 2001:240). Blust (1999b) said PAN *R became *x* in Pazih. This is so word-finally; however, the word-initial segment is unexpectedly *r*. The initial consonant seems to dissimilate to avoid two identical consonants as in the case in Proto-Atayal “leaf,” in which *gabag became *abag. In the case of Pazih, the connection of “leaf” to “summer” is much clearer than in the Atayalic subgroup. In Li and Tsuchida (2001:240), “leaf” is *rabax* and “summer” is *rabaxan*. This form also means “south.” *Rabaxan* “summer, south” is derived from *rabax* “leaf” through the suffixation of *-an*, as in the Atayalic subgroup.

2.3 Taokas

“Leaf” appears in a non-cognate form, *bixax*. However, a candidate for the cognate, *rabax* or *raba*, is observed under the item for “tea” (Tsuchida 1982:75). The semantic connection is so evident that it is possible to predict the meaning shift from “leaf” to “tea.” Regarding the sound correspondence, Ross (2015) notes that PAN *R is reflected as *h* in Taokas. Then, the expected form is *habah*. In the actual forms *rabax* and *raba*, the final consonant, *x*, is proximal to this segment. A possibility is that the transcription at the time of the interview was not sufficiently accurate, and *h* was written as *x*. The segment *x* is further lost in the other variation. Nonetheless, the initial consonant, *r*, is a problem; however, it is inferable that dissimilation from *h* to *r* operated analogous to the Pazih example (the expected form is *xabax* but the actual

form is *rabax*). Another possibility is that Taokas borrowed this word from Pazih from *rabax*.

2.4 Siraya

“Leaf” is transcribed as *hapa* (Adelaar 2011) or *haba* (Bullock 1874), which are highly likely to be the reflex of PAN *RabaR. Adelaar’s form, based on Bible translation written in the 17th century by Dutch missionaries (Gravius 1661, 1662), shows the devoicing of the medial consonant *b*. Regarding the reflex of PAN *R, Adelaar observed that it becomes *h*, *x*, or zero. Then, *hapa* or *haba* shows that the regular sound changes. The initial *R became *h*, and the final *R became zero.

2.5 Thao

“Leaf” appears in a non-cognate form, *fiŋaŋ*, in Blust (2003). However, Ogawa (2006) listed a form similar to PAN *RabaR. He transcribed this form as *θabaš* with the meaning “dead leaf.” Blust (2003:76) said PAN *b is reflected as *f* in Thao; however, at least in this example, the segment is retained as *b*. More problematic are the initial and final consonants. Blust (2003:79) said PAN *R is reflected as *l* in Thao. Then, the expected form is *lafal*. Ogawa’s initial and final consonants, *θ* and *š* (the latter is probably *f*), are similar to *l* in terms of the frication. It is probable that one of them was *l* but was transcribed slightly differently by Ogawa, and the other segment underwent dissimilation so that the initial and final segments differed. In the aforementioned examples of dissimilation (Atayal, Pazih, and Taokas), the initial segment underwent dissimilation. By analogy, in Thao, the initial segment might show the result of dissimilation and the final segment is the regular reflex. Then, Ogawa’s form can be re-transcribed as *θabal*.

2.6 Saisiyat

“Leaf” appears in a non-cognate form, *biya*, in Ogawa (2006). Under the item “leaf,” there seems to be no cognate. However, the cognate can be observed in “summer” as *haba:an* (Council of Indigenous Peoples 2017). This word must be derived by adding the suffix *-an* to a root, which is *haba* (the length of the vowel is disregarded here).

Ross (2015:32) indicated that PAN *R is reflected as *l* in Saisiyat. Then, the expected form is *labaŋ*. In *haba*, the form obtained through analysis, the initial consonant is *h*, which is voiceless such as the *l* is. The final consonant seems to be

lost. A probable explanation is that the segment *ʃ* further underwent a change to *h* so that *ʃabaʃ* became *habah*. Afterward, the final *h* was lost. The loss of the final consonant is hinted by the compensatory lengthening of a vowel in the form attached by *-an*.

The cognate is observed in “summer” and in “south.” Utsushikawa (1936) reported that “south” in Saisiyat is *kap-na-abaan*. The sequence in point is *abaan*; the remainder, *kap-na*, seems to be a type of prefix indicating a direction. This form is said to be identical to *haba:an* “summer,” although they differ slightly in that *abaan* has no initial consonant and no vowel lengthening. Saisiyat derived “summer” and “south” from *haba(h)* “leaf,” but the latter was lost afterward.

2.7 Bunun

“Leaf” appears in a non-cognate form, *lisav* (Ogawa and Asai 1935). Under the meaning of “leaf,” no cognate was observed. However, similar to Saisiyat, the form for “summer” seems to contain the cognate. “Summer” is *talabal* (Council of Indigenous Peoples 2017). This form is synchronically unanalyzable; however, it could be composed of a prefix *ta-* and a root *labal*. The hypothesized root form shows a reasonable reflex from PAN *RabaR. Li (1988) said that PAN *R becomes *l* in Bunun. Therefore, the sound correspondence is perfect. In addition, some languages, namely, Atayal, Seediq, Pazih, and Saisiyat, have derived “summer” from “leaf.” Thus, it is probable that *talabal* is derived from an earlier form for “leaf,” *labal*; however, this form and meaning were lost, and the origin of “summer” became opaque. However, Bunun is different from other languages in its derivational process. Other languages have the suffix *-an*, whereas Bunun has the prefix *ta-* (Table 5).

2.8 Summary of *RabaR

Table 5 summarizes the findings thus far. From Proto-Atayalic to Siraya, *RabaR means “leaf.” It means “dead leaf” and “tea” in Taokas. The root itself is lost in Saisiyat and Bunun but retained in a derived form meaning “summer.”

Table 5: *RabaR in Formosan languages

	Leaf	Summer	South
Proto-Atayalic (Proto-Atayal/ Proto-Seediq)	*Rabag (*abag/*rabag)	*Rabag-an (*gəbag-an/*rəbag-an)	
Pazih	<i>rabax</i>	<i>rabax-an</i>	<i>rabax-an</i>
Taokas	<i>rabax</i> “tea”		
Siraya	<i>hapal/haba</i>		
Thao	<i>θabal</i> “dead leaf”		
Saisiyat		<i>haba:-an</i>	<i>aba-an</i>
Bunun		<i>ta-labal</i>	
PAN	*RabaR		

3. Supplement: Proto-Austronesian “leaf” and “Alocasia”

In relation to the PAN reconstruction of “leaf,” the reconstruction of PAN “Alocasia” is discussed in passing. PAN *biRaq was reconstructed in Blust and Trussel (2010). According to them, this form means “an inedible tuber with large leaves: wild taro, elephant’s ear or itching taro, *Alocasia* spp.” For their reconstruction of *biRaq “wild taro,” they relied on three Formosan languages: Amis, Rukai, and Paiwan (Table 6). Here, PMP indicates a reconstructed Proto-Malayo-Polynesian form.

Table 6: Formosan data for *biRaq in Blust and Trussel (2010)

Amis	<i>fila?</i>	“leaf used to wrap betel nut for chewing; plant sp.”
Rukai	<i>viɔa, biɔa</i>	“wild taro: <i>Alocasia macrorrhiza</i> ”
Paiwan	<i>viaq</i>	“leaf used in rites (may be of various plants)”
PMP	*biRaq	“wild taro”
PAN	*biRaq	“wild taro”

They noted as follows in the footnote:

*“The known Formosan evidence for *Ribaq in the meaning ‘Alocasia macrorrhiza’ is confined to two dialects of Rukai. Far better attested in Taiwan are reflexes of a form with the same shape meaning ‘leaf’ ... If we are dealing with a single cognate*

set, then, the original specific reference to Alocasia appears to have been all but completely lost in Formosan languages...

Table 7: A cognate set for Formosan “leaf”

Language	Leaf ²	Alocasia ³
Atayal (Squliq)	<i>abaw</i>	<i>bəgayaw</i>
Seediq (Paran)	<i>wasaw</i>	<i>burayo</i>
Pazih	<i>rabax</i>	<i>biarax</i> ⁴
Bunun	<i>lisav</i> ⁵	<i>baial, baihal</i> ⁶
Saisiyat	<i>biya</i>	<i>byaraL</i>
Thao ⁷	<i>fiłaq</i>	<i>farazay</i>
Rukai	<i>biya</i>	<i>thiagenge</i> ⁸ , <i>biʔa</i>
Paiwan	<i>vīaq</i> “leaf used in rites”	<i>qayŋuay</i> ⁹

² The data for Saisiyat, Rukai, Puyuma, Basay, Taokas and Babuza are taken from Ogawa (2006:402). For Basay, Tsuchida et al. (1991) recorded another from *bela* “leaf.”

³ The data for Saisiyat *byaraL*, Rukai *biʔa*, Tsou, Kavalan and Kanakanavu are taken from Li (1994:786).

⁴ The data are taken from Li and Tsuchida (2001:89).

⁵ For Bunun, *vila* “piper betle” is recorded in Hu (2016:358), which may be a reflex of PAN *biRaq. However, I could not elicit a word for “piper betle” from my Bunun informant (Takibakha dialect) in Nantou County. This leaf is foreign to the Bunun tribe because they did not have the habit of chewing betel nuts in the old days, as my informant explained. I assume that this form recorded in Hu represents a loanword from Amis. This Bunun dictionary by Hu concerns the Bunun dialect in Taitung County in southeastern Taiwan. The Bunun villages in Taitung are adjacent to Amis villages. These Bunun tribes migrated to southeastern Taiwan from Nantou County in central Taiwan. In the course of migration, they must have adopted certain customs and words of the neighboring tribes. Additionally, in Bunun, the expected reflex of PAN *biRaq would be *bilaq* or *bilah* depending on the dialects. The attested Bunun form *vila* is similar to Amis *fiłaʔ* in that their initial consonants are fricatives.

⁶ The data are taken from Hu (2016:18). In his Bunun dictionary, Nihira (1988:30) recorded this word as *baial* or *baihal*. The second form has a medial *h*. For Takibakha Bunun, I elicited *behal* (< earlier form *baihal*), which also has the medial *h*.

⁷ Thao data are taken from Blust (2003).

⁸ The data are taken from Council of Indigenous Peoples (2017).

⁹ The data are taken from Ho (1978:612).

Puyuma	<i>bira</i>	
Kavalan	<i>bira</i> ¹⁰ “piper betle”	
Amis	<i>filaʔ</i> “piper betle”	<i>tabayal</i> ¹¹ (Sakizaya dialect)
Basay	<i>beya</i>	
Taokas	<i>bixax</i>	
Babuza	<i>bia</i>	
Tsou	<i>ehujɯ</i> ¹²	<i>cohu</i>
Kanakanavu	<i>ranuɯɯj</i> ¹³	<i>caunu</i>
Saaroa	<i>rahlujɯ</i>	<i>caluʔu</i>

The cognate set for the Formosan “leaf” that they have in mind is listed in Table 7 with the shade including the cognate in Rukai, which also means “leaf.” Table 7 also shows the forms for “Alocasia” in Formosan languages. If PAN *biRaq means “Alocasia,” it does not only mean that almost all Formosan languages have shifted its meaning to “leaf” but also that the words for “Alocasia” were replaced by other words listed in Table 7 because almost all Formosan languages have distinct forms for “Alocasia.”¹⁴

The data in Table 7 show that all the attested forms of *biRaq (the shaded forms) in Formosan languages have the basic meaning of “leaf” and that the words for “Alocasia” are different from *biRaq except for Rukai *biya/biʔa*, which seems to mean both “leaf” and “Alocasia.”

Regarding “Alocasia,” Cauquelin (2015:429) noted that the Puyuma word *siadeng* is a Southern Hokkien loanword. It is likely that Rukai has the cognate *thiagenge* also borrowed from Southern Hokkien. Rukai has another form *biʔa*; however, its cognates in other Formosan languages mean “leaf” (Table 7). Tsou, Kanakanavu, and Saaroa, belonging to Tsouic subgroup, demonstrate they have other cognate sets.¹⁵ Paiwan demonstrates yet another cognate.

¹⁰ The data are taken from Li and Tsuchida (2006:87).

¹¹ The data are taken from Council of Indigenous Peoples (2017).

¹² The data are taken from Nevsky (1993:68).

¹³ The data for “leaf” in Kanakanavu and Saaroa are taken from the Council of Indigenous Peoples (2017).

¹⁴ Data for Squiliq Atayal, Paran Seediq, and Bunun are collected by the author unless cited otherwise.

¹⁵ Li (1994:786) reconstructed Proto-Tsouic *caluʔu.

In Li's (1994) list for Formosan "Alocasia," he classified Atayal and Seediq (his data were Atayal *bagayaw*, *bgayax* and Seediq *bgayaw*) into one group and Saisiyat and Pazih into another group, reconstructing Proto-Atayalic and Proto-Saisiyat-Pazih *bagayag and *byaraR, respectively.

Further, Atayalic could share the same cognate set with Bunun and Sakizaya Amis. The tentative reconstruction for this group would be PAN *bayaR. Of the reflexes of this tentative proto-form, only Bunun *baial* retains its original form (However, a variant *baihal* is problematic in that it has the medial *h* that does not exist in Atayalic and Sakizaya Amis). Sakizaya Amis *tabayal* added the prefix *ta-*. For Atayalic, the tentative proto-form is *b<aR>ayaR with an infix <aR>. This form is reflected as *b<ag>ayaw* (Li 1994) in Pngawan Atayal, *b<əg>ayax* (Li 1994) in Skikun Atayal, and *b<əg>ayaw* in Squliq Atayal. This form is reflected as *b<ər>ayaw* in Truku Seediq (Rakaw 2006:138) and *b<ur>ayo* in Paran Seediq.¹⁶

4. Conclusion

This paper reconstructed PAN "leaf" as *RabaR. In the context of Formosan languages, three proto-forms mean "leaf": *waSaw, *biRaq, and *RabaR. Table 8 summarizes the findings on the proto-form of "leaf." Each proto-form of "leaf" is listed along with the language(s) and corresponding reflexes (the reflexes for *RabaR show reconstructed forms in Bunun, Atayal, Seediq, Thao, and Saisiyat; the reconstructed forms in Bunun and Saisiyat are no longer used independently).¹⁷

¹⁶ Seediq form is taken from the author's field notes. This infix pattern is also likely to be observed in Thao *f<ar>azay*; however, if we suppose it belongs to the same cognate set as Atayalic, Bunun, and Sakizaya Amis (*bayaR), it is difficult to explain the final consonant *y* because the PAN *R should be reflected as *l* in Thao. It is also difficult to explain the medial consonant *z*, which is a reflex of PAN *j, not *y. Furthermore, if we suppose Saisiyat and Pazih also belong to the same cognate set, an infix <i> seems to be added, i.e., *b<i>ayaR; however, it is also difficult to explain the medial *r* in Pazih *biarax* and Saisiyat *byaraL*. The expected segment here is *y*.

¹⁷ No data for "leaf" were found in Ketagalan and Hoanya.

Table 8: Three “leaves” in Formosan languages

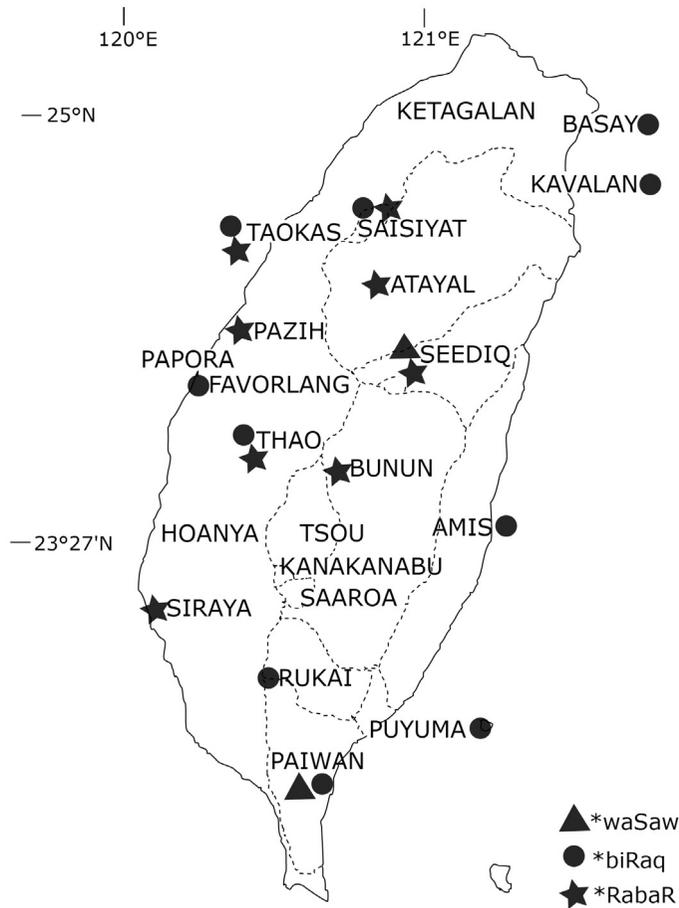
*waSaw	*biRaq	*RabaR
		Siraya <i>hapa/haba</i>
		Pazih <i>rabax</i>
		Bunun <i>*labal</i>
		Atayal <i>*abag</i>
Seediq <i>wasaw</i>		Seediq <i>*rabag</i>
	Thao <i>filaq</i>	Thao <i>*θ abal</i>
	Taokas <i>bixax</i>	Taokas <i>rabax</i>
	Saisiyat <i>biya</i>	Saisiyat <i>*haba(h)</i>
Paiwan <i>asaw</i>	Paiwan <i>viaq</i>	
	Puyuma <i>bira</i>	
	Rukai <i>biya</i>	
	Basay <i>beya</i>	
	Kavalan <i>bira</i>	
	Babuza <i>bia</i>	
	Amis <i>fila?</i>	

Having three proto-forms for “leaf” at the same time could be unreasonable. There could be a chronological order for the words. However, the author is unable to make any diachronic inference at the proto-language level.

Map 1 shows the geographical distribution of the three reconstructed forms for “leaf.” This geographical distribution does not show clear-cut areal features. For *waSaw, Seediq and Paiwan are discontinuous. Regarding *biRaq, there are three discontinuous groups: (i) Taokas, Saisiyat, Favorlang, and Thao; (ii) Paiwan, Puyuma, Rukai, and Amis; and (iii) Basay and Kavalan. For *RabaR, all the languages are more or less contiguous. The geographical distribution is too interlacing to draw a conclusion regarding the antiquity of these forms. To further complicate the situation, some languages have two proto-forms. In addition, the combination of the two proto-forms differs among them: Seediq has *waSaw and *RabaR; Thao, Taokas, and Saisiyat have *biRaq and *RabaR; and Paiwan has *waSaw and *biRaq. At this stage, the only argument that can be made is that three

“leaf” forms are reconstructed.

Map 1: Distribution of Formosan “leaves”



If there were any differences among the three forms, they could have been semantic. One may be a general term, and others may refer to a special type of leaf. With regard to semantics, *RabaR is peculiar in that it developed the meaning “summer” in five languages (Atayal, Seediq, Saisiyat, Pazih, and Bunun), which could literally mean “leaf season,” or it even developed the meaning “south” in two languages (Saisiyat and Pazih), which could semantically be related to the hotness of summer. In terms of the form, “summer” or “south” is derived by affixation. This formal change and the semantic shift indicate the archaic feature of *RabaR compared with the other two, which seem not to have these changes.

Finally, the connection with non-Formosan languages is overviewed.

According to Blust and Trussel (2010), *waSaw “leaf” is observed in Formosan languages only; by contrast, the reflexes of *biRaq meaning “leaf” are widely observed in Formosan languages, and its reflexes meaning “Alocasia” are widely observed in Malayo-Polynesian languages. In addition, they posit another form for “leaf” in Proto-Malayo-Polynesian, *dahun.¹⁸

According to the author’s review of the literature, *RabaR was observed in Formosan languages only. However, a cognate could be found in non-Formosan Austronesian languages with forms and meanings that deviated from the original ones, because it barely survived in Bunun and Saisiyat in the derived form meaning “summer” or “south.”

References

- Adelaar, Alexander (2011), *Siraya: Retrieving the phonology, grammar and lexicon of a dormant Formosan language*. Berlin: De Gruyter Mouton.
- Blust, Robert (1999a), Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*, edited by Elizabeth Zeitoun and Paul Jen-kuei Li, 31–94. Taipei: Institute of Linguistics, Academia Sinica.
- Blust, Robert (1999b), Notes on Pazeh phonology and morphology. *Oceanic Linguistics* 38: 321-365.
- Blust, Robert (2003), *Thao dictionary*. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
- Blust, Robert and Stephen Trussel (2010), *Austronesian comparative dictionary, web edition*. <http://www.trussel2.com/ACD/> [accessed on Apr 9].
- Bullock, Thomas L. (1874), Formosan dialects and their connection with the Malay. *China review, or notes and queries on the Far East* 3:38-46.
- Cauquelin, Josiane. (2015), *Nanwang Puyuma-English dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Council of Indigenous Peoples (2017) *Online Formosan dictionary* <http://e-dictionary.apc.gov.tw> [accessed on Apr 9].

¹⁸ Nevsky (1993:68) says Tsou *ehujju* “leaf” is a cognate with Malayo-Polynesian reflexes of *dahun. However, Sagart (2014:876) says Proto-Tsouic *rałəŋa “leaf,” which was reconstructed by Tsuchida (1975:326), is an innovative word in Proto-Tsouic and it has no cognates outside Tsouic subgroup.

- Ferrell, Raleigh (1969), *Taiwan aboriginal groups: problems in cultural and linguistic classification*. Monograph No.17, Institute of Ethnology, Academia Sinica. Taipei: Academia Sinica.
- Gravius, Daniel (1661), *Het heylige Euangelium Matthei en Johannis. Ofte Hagnau ka d'llig matiktik ka na sasoulat ti Mattheus ti Johannes appa*. Amsterdam: Michiel Hartogh.
- Gravius, Daniel (1662), *Patar ki tna-'msing-an ki Christang ofte 't Formulier des Christendoms*. Amsterdam: Michiel Hartogh.
- 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 (1988), A comparative study of Bunun dialects. *Bulletin of the Institute of History and Philology, Academia Sinica*, 59(2):479–508.
- Li, Paul Jen-kuei (1994), *Some plant names in Formosan languages*. In *Selected Papers on Formosan Languages*, 767–804. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei and Shigeru Tsuchida (2001), *Pazih dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Li, Paul Jen-kuei, and Shigeru Tsuchida (2006), *Kavalan dictionary*. Institute of Linguistics, Academia Sinica.
- Ho, Dah-An (1978), A comparative study of five Paiwan dialects. *Bulletin of the Institute of History and Philology* 49.4:565–681.
- Hu Jin-sheng, ed. (2006), *Bunong zuyu jidian*. Taitung: Taoyuan Elementary School.
- Nevsky, Nikolai A. (1993). *Taiwan zouyu yudian*. Translated by Bai Si-hong, Li Fuqing, and Pu Zhong-cheng. Taipei: Taiyuan.
- Nihira, Yoshiro (1988), *A Bunun vocabulary: A language of Formosa*. Tokyo: Ado-in Kabushiki Kaisha. Third edition.
- Ogawa, Naoyoshi (2006), *A Comparative vocabulary of Formosan languages and dialects*, compiled by Paul Jen-Kuei Li and Masayuki Toyoshima. Tokyo: Research Institute for Languages and Cultures in Asia and Africa.
- Ogawa, Naoyoshi and Asai Erin (1935), *The myths and traditions of the Formosan native tribes*. Taipei: Taihoku Imperial University.
- Pecoraro, Ferdinando (1979), *Eléments de grammaire Taroko, précédés de la présentation de la vie et de la culture des Taroko*. Paris: Association Archipel.
- Rakaw, Lowsi, Jiru Haruq, Yudaw Dangaw, Yuki Lowsing, Tudaw Pisaw, and Iyuq Ciyang eds. (2006), *Tailugezuyu jianyi zidian*. Hualian: Xiulin Township

Administrative Office. [In Mandarin]

- 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.
- Sagart, Laurent (2014), In defense of the nurnearl-based model of Austronesian phylogeny, and of Tsouic. *Language and Linguistics* 15(6):859–882.
- Tsuchida, Shigeru (1975), Reconstruction of Proto-Tsouic phonology. Ph. D. diss, Yale University.
- Tsuchida, Shigeru (1982), A comparative vocabulary of Austronesian languages of sinicized ethnic groups in Taiwan, part I. West Taiwan. *Memoirs of the Faculty of Letters, University of Tokyo*, No.7. Tokyo: Faculty of Letters, University of Tokyo.
- Tsuchida, Shigeru, Yukihiro Yamada, and Tsunekazu Moriguchi (1991) *Linguistic materials of the Formosan sinicized populations I: Siraya and Basai*. Tokyo: The University of Tokyo, Linguistics Department.
- Utsushikawa, Nenozo (1936), Time-conception in the primitive society. *The Japanese journal of ethnology* 2(4): 775–803.

Acknowledgments

This research is funded by JSPS KAKENHI Grant number JP1700060. The author presented an earlier version of this paper at the Tibeto-Burman Linguistic Circle on April 21, 2018. I am grateful to the participants who gave me comments on this occasion. I was also benefitted from two anonymous reviewers and Syuntarô Tida’s comments. However, any remaining faults belong to the author.

Formosan “leaf”: A reconstruction

Izumi OCHIAI

Kobe City University of Foreign Studies

ABSTRACT

This paper reconstructs the proto-Austronesian word for “leaf” as *RabaR on the basis of the forms for the said word in the following Formosan languages: Atayalic languages (Atayal and Seediq); Pazih; Siraya; and Thao. Additionally, forms for “tea” in Taokas and for “summer” in Saisiyat and Bunun are included as cognates. It was found that in the context of Formosan languages, there are three proto-forms for “leaf,” including *waSaw and *biRaq, that have already been reconstructed.

Keywords: Formosan languages, Atayalic, Historical linguistics, leaf, Alocasia