

New Data on Three Extinct Formosan Languages

Paul Jen-Kuei Li

Abstract: Some field data is presented for three extinct Formosan languages, namely: Basay, Siraya and Taokas, and some conclusions are drawn as to their linguistic value and historical derivations from Proto-Austronesian (PAN) or Proto-Hesperonesian-Formosan (PHF). Through this paper, I hope to stimulate more research concerning these soon to be lost languages, since all Formosan languages are threatened with extinction today. Some suggestions for preserving these languages are also given in this paper, including the recently devised orthographic systems.

1. Introduction

Among about twenty Formosan languages, some have long become extinct, while a few others are on the verge of extinction.¹ The extinct languages are all "plain tribe" languages, including Ketangalan and Basay in the northernmost part of Taiwan, and Taokas, Babuza, Papora, Hoanya and Siraya in the western part of Taiwan. The nearly extinct languages include not only the "plain tribes", such as Kavalan, who originally settled in the northeastern part of Taiwan, and Pazeh, who settled in the central western part of Taiwan, but also some of the so-called "mountain tribes", Thao, Saisiyat, Kanakanavu and

1 An earlier version of this paper was presented to the Symposium on Austronesian Language Survival Issues, the Sixth International Conference on Austronesian Linguistics, Honolulu, May 20-24, 1991. I wish to acknowledge the editorial assistance of Kathleen Ahrens and her valuable suggestions to improve an earlier draft of the paper. I am also indebted to Paul Benedict, Dah-an Ho and Shigeru Tsuchida for their helpful comments.

Saaroa. Some field work has been done on these nearly extinct languages: Kavalan by Li (1978, 1982), Moriguchi (1982, 1983), and Tsuchida, Pazeh by Tsuchida (1969), Ferrell (1970), Li (1978), Li and Lin (1990), Thao by F. K. Li et al (1959) and Li (1976, 1978), and so on.

To a greater or smaller extent, all Formosan languages are threatened with extinction today. They are all minority languages, with a total population of no more than 400,000 out of the twenty million people in Taiwan. They are all rapidly losing their cultural and linguistic identities as the pace of sinicization hastens. To make matters worse, most of them have no writing systems of their own even today, and there are few written documents for most of the languages, especially the ones that are extinct. Records by western missionaries and Japanese scholars before World War II have left remnants of data for the extinct languages. Aside from a few western and Japanese scholars, only a few Chinese scholars and students have worked on Formosan languages since 1945.

In this paper, I will present some new data for three extinct Formosan languages, namely: Basay, Siraya and Taokas, and draw some conclusions as to their linguistic value and historical derivations from PAN or PHF. In this way, I hope to stimulate more research concerning these soon to be lost languages.

2. Studies of the Plain Tribe Languages

By the time Japan first occupied Taiwan in 1895, most of the plain tribes had already lost most of their own languages and customs. In their monumental work on Formosan languages, Ogawa and Asai (1935:4) simply mentioned the state of the plain tribes, but presented no data for any of these languages. Fortunately Dutch and Chinese documents written before the occupation had left behind some information for these languages. Japanese scholars, especially Ino, Ogawa and Asai, did make efforts and take pains to record languages of the plain tribes that were still remembered by a few old people in various villages and locations primarily in the first half of the 20th century. Their collections were

mostly wordlists and some phrases for languages of various plain tribes. Some of the data was made available in the form of comparative vocabularies in Tsuchida's (1982a, 1985) and in Tsuchida *et al*'s (1991) recent publications. Tsuchida (1982a) contains some 400 lexical items for Taokas, Babuza, Papora, Hoanya, Pazeh and Siraya, while Tsuchida (1985) contains only less than 100 items for Basay, Ketangalan and Kulon. Tsuchida *et al* (1991) provides 163 lexical items for 3 major Siraya dialects originally prepared by Ogawa and hundreds of lexical items for Basay and Trobiawan originally collected by Asai.

It is hard to say exactly when a language becomes extinct. Field workers can occasionally extract a few words, or even dozens of words, from older speakers here and there for languages that have long been considered extinct. Tsuchida (1982b), for instance, was able to collect about 80 words and phrases from a 71 year-old Papora speaker late in 1969. Even in 1987, I was still able to collect 67 words and phrases for Basay from an 83 year-old woman presently living in Jilung.

We have imperfect understanding for the phonetic details of all these extinct languages. None of the informants these days have a native command of the languages. Their pronunciation of the Formosan forms may have been influenced by Taiwanese, a native Southern Min dialect of Chinese, which they speak natively.

2.1 Basay

Scholars such as Abe (1930) and Ogawa (1944) used to refer to Basay as Ketangalan until Mabuchi's (1953) amendment; see Tsuchida (1985).

Basay was formerly spoken in the plains of Taipei and Jilung, and partly in the northeastern coastal areas of Ilan, where Kavalan was formerly spoken. Aside from Tsuchida *et al*'s (1991) Basay data originally collected by Erin Asai in 1936, very little linguistic information about Basay is available in print, as there is little data that has been recorded in the past. There are no native speakers alive today.

My recording of the following Basay is based on an old woman named Pan-rau

Tzeng (曾潘蟻), who originally came from Shinshe, Gungliau (貢寮新社), where she picked up the following words in her childhood from an old Basay woman with gray hair.² Stress falls on the final syllable. The stops and the affricate /c/ [ts] are generally unaspirate. The Kavalan forms cited after a few Basay forms for comparison below are also based on my own field notes.

1. tina 'mother' < PAN *(t)ina
2. tama 'father' < PAN *(t)amaH
3. baki 'old man or woman' < PAN *baki 'grandfather'
4. pakuayan 'elder brother'
5. binay haliuhiu 'a woman with improper behavior'
 b-in-ay 'woman' < PAN *beHi
6. torang 'a man with illicit sexual relations'
7. basay 'natives, aborigines, self-appellation'
8. bucut 'Chinese'
9. cima 'hand' < PAN *lima
10. mata 'eye' < PAN *maCa
11. bukuce³ 'hair' < PAN *buSek
12. tuku (metathesis of consonants) 'head louse' < PAN *kuCu
13. boolo 'ear'
14. nganga 'mouth'
15. bangcaw 'tooth', cf. Kavalan vangRaw
16. cacace 'tongue'
17. cicu 'breasts' < PAN *t'ut'u
18. kulupu 'eyelash', cf. Kavalan qLupu

2 I am grateful to Dung-nan Wu (吳東南) for informing me of the informant and taking me to Jilung to record the language.

3 Note the metathesis of *k and *S, a sound change shared by either Formosan or extra-Formosan languages. Since the informant is not a native speaker of Basay, she may have added vowel [e] after the final consonant /c/[ts]. See also Nos.16, 23 and 37 below.

19. babuy 'pig' < PAN *babuy
20. wacu 'dog' < PAN *wat'u
21. tarahok 'chicken' < Proto-Formosan (PFN) *teRakuk⁴
22. kalabaw 'carabaw' cf. kalabaw or karabaw in some
Philippine languages (see Reid 1971:59)
23. bulace 'husked rice' < PAN *beRat'
24. sumay 'cooked rice' < PAN *Sem[ae]y
25. tabun 'side dish', cf. Kavalan tamun
26. hawpit 'sweet potato', cf. Kavalan qawpiR
27. ʔut 'sugarcane'
28. tabaku 'tobacco'
29. takil 'match'
30. ciabang 'bowl' (loan from Japanese)
31. apiat 'chopsticks'
32. sulukən 'clothes'
33. pila 'money' < PAN *pirak
34. lang 'wine', keman lang 'to drink wine'
35. solongsong 'type of seaweed'
36. k-əm-an 'to eat' < PAN *kaʔen
37. t-om-angice 'to weep' < PAN *Cangit'
38. acap 'to take' < PAN *alap
39. rakaw 'to sleep'
40. sa-(ka-)kan 'to cook' < PAN *kaʔen
41. bolo-bolo 'to bathe' (with /l/ rather than the anticipated /c/) < PAN *bulu

4 Cf. Tso troʔu-a, Kan tarikuuk-a, Sar turukuuk-a, RukTa tarukuku, Bun tulkuk, Puy terkuk, Kav tLaquq, Ami tuLakuk 'chicken'. (Abbreviations for Formosan language names are: Tso, Tsou; Kan, Kananav; Sar, Saaroa; RukTa, Rukai (Tanan dialect); Bun, Bunun; Puy, Puyuma; Kav, Kavalan; Ami, Amis). The h in the Basay form is irregular; the expected reflex is k.

42. bolo kaba 'to wash clothes', cf. 75 below
43. parubi 'to gamble'
44. kualaw na 'lost, invisible'
45. patay 'to hit others' < PAN *paCey
46. mamanaa 'to keep quiet'
47. kualaw cima 'to have no hand'
 kualaw 'none' < PAN *waDa
48. kimikimi 'unable to see'
49. moosakusa 'Don't talk'
50. ʔawi 'to go, to return'
51. mu-ali 'to come' < PHF *[hS]ali, cf. Pazeh ali, Tagalog hali ka 'Come! (imp.)'
52. napat 'to defecate'
53. s-om-ilaw 'hungry' < Proto-Southern-Formosan (PSF) *siLaw⁵
54. sangangiauw 'pretty'
55. hahuy 'bad'
56. unuwa sayan 'What are you doing?'
57. siarui sialai 'to go back and forth, to idle'
58. maralaw 'to die'
59. c'a⁶ 'one' < PAN *etʔa
60. lusa 'two' < PAN *Dusa
61. c'uu 'three' < PAN *teluH
62. sipat 'four' < PAN *Sepat
63. anəm 'six' < PAN *enem⁷
64. pitu 'seven' < PAN *pitu
65. wacu 'eight' < PAN *walu

5 Cf. Tso siro (srov-a), Ruk ma-siLaw 'hungry'.

6 The affricate is aspirate, indicated by the symbol [ʔ], only in a few forms.

7 My informant mixed up the numerals and did not give the word for 'five'; see No.111 below.

- 66. siwa 'nine' < PAN *Siwa
- 67. labatan 'ten'

The following Basay data was based on Tsuchida (1985), which was originally collected by previous Japanese scholars including Ino, Ogawa and Asai:

- 68. ucu 'head' < PHF *quluH
- 69. ngucucu 'nose'
- 70. hai 'foot, leg'
- 71. lais 'face' < PAN *daqis
- 72. cala 'blood' < PAN *DaRaQ
- 73. ciku 'elbow' < PAN *t'iku
- 74. cubu (<A⁸ i>u) 'urine' < PHF *t'ibu
- 75. kaba 'jacket, coat, clothes' < PFN *kaba⁹
- 76. aalat 'fence' < PAN *qalad
- 77. kinien 'house'
- 78. rayar (<A c->r-) 'sail' < PAN *layaR, cf. Paiwan la-laya 'flag' (Ferrell 1982:156)
- 79. vocal (-a- irregular) 'bow' < PAN *but'uR
- 80. ramar 'fire' < PAN *DamaR, cf. Kavalan zamaR 'fire'
- 81. abu 'ash' < PAN *qabu
- 82. tau 'person' < PAN *Cau
- 83. vanak 'child' < PAN *aNak
- 84. kaka 'elder sibling' < PAN *kaka
- 85. vaki 'grandfather' < PAN *baki
- 86. vai 'grandmother' < PAN *beHi
- 87. cawaa 'husband' < PAN *t'awa
- 88. cenal 'sun' < PHF *t'iNaR

8 Hereafter A indicates assimilation and M metathesis.

9 Cf. Paiwan kava 'clothes, snake slough'.

89. uran 'rain' < PAN *quDaN
 90. lanum 'water' < PAN *DaNum
 91. baci 'wind' < PFN *bali
 92. lesau 'tree'
 93. bucawan 'gold' < PHF *bulawan, cf. Kavalan vuLawan
 94. cacan (<A r>c) 'path' < PAN *Zalan
 95. laya 'west' < PAN *Daya
 96. rabi, (lave in Abe 1930:430) 'night' < PAN *Rabi?iH
 97. vautte, (vaut in Abe 1930:431) 'fish', cf. Kavalan vaut
 98. kulaba 'duck', cf. Kavalan kLava
 99. telo(t) 'egg' < PAN *qiCeluR
 100. rotong 'monkey' < PHF *luCung
 101. timula, tomusa 'flea' < PSF *qatimula
 102. sarak 'good'
 103. law 'big'
 104. ca-cawe 'far' < PAN *d'aweq
 105. mu-za-zani 'near' < PHF *[dD]aNih
 106. zazan 'old' < PHF *[dD]a[dD]aN
 107. raran 'thousand' < *[dD]a[dD]an
 108. racibu (<M) 'hundred' < *t'a-ribu, cf. also Kavalan mRasivu (<M)
 109. baci u 'to sell' < PFN *baliw
 110. rareu 'to run away' < PAN *laRiw
 111. c-um-ulat 'to write' < PHF *t'uNat
- The following Basay ("Ketangalan") data is based on Abe (1930:430-435):
112. cima 'five' < PAN *lima
 113. k-an-uukus 'claw', cf. Kavalan q-n-uuqus < PAN *ku(S)kuS
 114. luna, zauzau 'ants'

The following Basay ("Ketangalan") data is based on Ogawa (1939, 1943):

- 115. Lusu? 'tears' < PAN *luSeq
- 116. kenava 'mortar'
- 117. Li-cu 'pestle' < PHF *qaSeluH
- 118. taasau 'year', cf. Kavalan ta-tasaw
- 119. bucan 'month, moon' < PAN *bulaN
- 120. li-en 'day', cf. Atayal(Squliq) qli-an 'daytime'
- 121. rabe-rabe 'morning', cf. Kavalan ta-Rav-Ravi
- 122. tobau 'yesterday'
- 123. wacia lien 'today'
- 124. ko-tobau 'tomorrow'
- 125. lasu-li-en 'day after tomorrow'
- 126. kinan 'when'

The following item is mentioned in Abe's (1938: 97) study of place names in Taiwan:

- 127. baŋka? 'canoe' < PAN *baŋka?

Even with the limited data above, we can get quite a bit of valuable information about the Basay language:

(1) We find dozens of PAN or PHF cognates in a few more than a hundred items, as indicated above. Some of them provide extremely valuable information. For instance, Nos. 78 and 127 indicate that Basay retained PAN cognates for both 'sail' and 'canoe', providing crucial evidence that Formosan peoples, like other Austronesian peoples, were also seafaring. See Li (1991a) for further discussion.

(2) Unlike most Formosan languages, Basay does not distinguish between PAN *t and *C, or *n and *N. Only a few Formosan languages merge *t and *C, *n and *N: Bunun, Amis,¹⁰ Kavalan, Siraya and Ketangalan (including Basay and Luilang). Ferrell (1969:69) failed to mention Ketangalan (now Basay) and misplaced Luilang in his "Paiwanic I" subgroup which distinguishes between *t and *C.

(3) PAN or PHF derivations to Basay are as follows: *p > p; *b > b or v; *t, *C > t; *d,

10 Although Amis and Siraya merge *t and *C, they distinguish between *n and *N.

Paul Jen-kuei Li

*D, *d', *R > l or r; *k > k; *q, *H, *ʔ > \emptyset ; *t', l > c; *S > s; *m > m; *n, *N, *g' > n; *ŋ > ŋ; *w > w; *y > y; *a > a; *i > i; *u > u; *e > ə, u, i. See examples above. There are a few exceptions to the above derivations.

(4) Several Basay forms closely resemble those in Kavalan; see Nos.15, 18, 25, 26, 93, 97, 98, 113, 118 and 121. These two languages were geographically adjacent to each other, and they may have commonly inherited these words or borrowed from each other.

2.2 Siraya

Siraya was formerly spoken in the southwestern coastal areas of Taiwan. It consisted of three subgroups: (1) Siraya Proper or Sidaia, (2) Taivoan and (3) Makatao, geographically running from north to south, and occupied fairly large geographical areas in west Taiwan.

Dutch missionaries such as Gravius (1661, 1662) provided biblical translations into the Siraya language, with corresponding versions in Dutch (and also English for Gravius 1661). Gravius (1661) was later edited by Campbell (1888) and Gravius (1662) by Asai (1939).

Murakami (1933) collected 116 Siraya manuscripts of land deeds, some of which have corresponding Chinese versions, and in the appendix he listed a Siraya vocabulary of approximately 1,000 items compiled from the Utrecht manuscripts, and gave both Dutch and English glosses.

In one of Ogawa's unpublished manuscripts, which was made available only recently in Tsuchida et al (1991), he lists 163 lexical items from various sources and locations to compare the different dialect groups of Siraya: 26 variants for Sidaia, 18 for Makatao and 26 for Taivoan. His data clearly indicate that the Siraya language falls into the three subgroups mentioned above.

All these publications and manuscripts provide us with invaluable data for the Siraya language.

I made two unsuccessful field studies of Siraya, initiated by false information about the availability of living speakers, first in Fuli county (富里鄉) in the east coast of

Taiwan in October 1976 and then in Yujing, Tainan (台南玉井) in the southern part of Taiwan in January 1981. I managed to collect only 15 lexical items from an 85 year-old man named Shi Wen (史温 whose grandmother spoke Siraya) from Muja (木柵) on January 11, 1981 as listed below.

1. kuka? 'chicken'
2. ?asu? 'dog' < PHF *wat'u
3. babuy 'pig' < PAN *babuy
4. luang 'water buffalo' < PHF *Nuang
5. kating 'child'
6. ?ugit 'bamboo broom'
7. tabuliang 'mouse'
8. tapungpung 'worms on bamboo shoots'
9. tabaku? 'tobacco, pipe'
10. kalalak 'husband of grand daughter'
11. ma?ilang 'to sleep'
12. ?u:u: 'Let's go home (to a child carried on one's back)'
13. mazangazang 'bad'
14. tabang 'preserved taro stem'
15. lupa? 'sausage with sticky rice as stuffing'

In addition, I elicited two terms (Nos.16 & 17), from a man in Toushe (頭社) and one term (No.18) from an old woman in Ganglin (岡林) :

16. ?abiki 'betel nut'
17. babuy 'pig' (same as No.3)
18. pawsin 'corn'

A man named Chuan-min Li (李全明 aged 62 in 1981) in his childhood learned the following numerals from his grandmother, who came from Sinkang (新港) and died at the age of 78 in 1930:

19. takabilum 'one'

Paul Jen-kuei Li

- 20. kulom 'three'
- 21. kulomta? 'four'
- 22. sakuba? 'five'
- 23. zangizang 'six'
- 24. kumsin 'seven'
- 25. tapatit 'nine'
- 26. kaytiang 'ten'

Some of the above 8 numerals do not look like any other Siraya dialect compared in Ogawa, or any other Austronesian language I know of. Only parts of the forms look similar to some of the numerals in some of the Makatao dialects, e.g. taka1 'one' in one dialect, kulun 'three' in another, takuba 'five' or 'six' and kunsin 'seven' in still another.

Aside from items Nos. 2, 3, 4, 16, 17 and above, all the other forms are not found in any other existing Siraya data although there are words with similar glosses.

The following sound system of Siraya is an approximation:

p	t	k	i	u
b~v	d~z	g~h		
	s		a	
m	n	ng		
	l			
	r			
w		y		

These are dialectal variants: b~v, d~z, and possibly also g~h.¹¹

Historical derivations from PAN or PHF to Siraya are the following. The Siraya forms are based on the Utrecht manuscripts cited in Murakami 1930 (I have reinterpreted the orthography in the Utrecht manuscripts: tt as t, ck as k, ch as c, ll as l, o or ou as u, ai as ai, etc.) as well as Ogawa [n.d.].

11 In traditional Dutch orthography, there is no contrast between g and h (K. A. Adelaar, p.c.).

*p > p	E.g. PAN *pitu > pitu 'seven' PHF *paCay > ma-patay 'to die'
*t > t	PAN *telu > turu 'three'
*C > t	PAN *maCa > mata 'eyes' PHF *Cawa > ma-tawa 'to laugh'
*k > k	PAN *kayu > kayu 'wood' PHF *baluku > baraku (A < a/u) 'winnowing basket' PHF *kaRang > kagan (<D n/ng) 'crab'
*b > b or v > f/_#	PAN *babuy > babuy 'pig' PHF *qeNeb > ilif 'to close'
*t' > s > ϕ /_#	PAN *wat'u > ?asu 'dog' PAN *et'ʔa > isa 'one' PAN *Cangit' > t-m-angi 'to cry'
*D > d, z or s	PAN *DemDem > ma-simdim 'dark' PHF *DapaN > zapan, sapat 'foot' PAN *DaNum > dalum, zalum 'water' PHF *Dilaq > da-dila 'tongue'
*d' > d or r	PAN *d'alan > darang, raran 'road'
*R > g or h	PAN *wiRiH > ugi 'left side' PHF *baRaH > vaga 'burning charcoal' PHF *baqeRuH > vahu 'new'
*S > g or h > ϕ /_#	PAN *DuSa > duha, zuha, 'two' PAN *Sepat > hpat 'four' PAN *paRiS > pagig 'rayfish' PHF *qamiS > ami 'north'
*q > ϕ	PAN *qabuH > avu 'ash' PAN *quDaN > udan 'rain'
*H > ϕ	PHF *qaSeluH > ha-ieru, ha-hilu 'pestle'

Paul Jen-kuei Li

*m > m	PHF *qayam > ayam 'bird' PAN *lima > rima 'five'
*n > n	PAN *enem > nnum 'six' PAN *wanaN > mu-nal 'right side'
*g' > n	PAN *[]ag'an > nanang (-ng irregular) 'name'
*ng > ng	PAN *CalingaH > tangira (<M) 'ear' PHF *qabang > avang 'boat'
*l > r	PAN *likuD > rikus 'back' PHF *luCung > rutung 'monkey'
*N > l (or n)	PHF *Nuang > luang 'carabaw' PAN *panDaN > pandal 'pineapple' PHF *aNak > alak 'child' PAN *bulaN > buran 'moon'
*w > v	PAN *Siwa > siva 'nine'
*y > y	PHF *qayam > ayam 'bird'
*i > i	PAN *(t)ina > ina 'mother'
*u > u	PAN *batu > batu 'stone' PAN kuCu > kutu 'louse'
*a > a	PAN *(t)ama > ama 'father' PAN *kaka > kaka 'elder sibling' PAN *Sapuy > apuy 'fire'
*e > i or u	PAN *et'ʔa > isa 'one' PHF *qeNeb > ilif 'to close' PHF *qaSeluH > ha-hilu 'pestle' PAN *enem > annum, nnum 'six'

Siraya shares only a very low percentage of cognates with other Formosan languages. Moreover, lexical evidence does not indicate a close genetic relationship with any particular

Formosan language (Li 1990). This 'unique' position of Siraya may be due to our present inadequate understanding of the language. If we make more careful studies of the existing biblical translations (Gravius 1661, 1662) and Sinkang manuscripts (Murakami 1933) mentioned above, we can certainly get a better understanding of the language. In fact, Siraya is one of the few extinct Formosan languages that have fairly extensive source materials to work with.

2.3 Taokas

I visited Shinmin village, Houlung (後龍新民) in July 1982,¹² and was able to collect the following words and phrases from several Taokas descendants there. The symbol ['] immediately after a stop stands for aspiration. Stress falls on the penult.

1. k'aʔan 'cooked rice' < PAN *kaʔen
 2. giakau 'wine'
 3. hipuʔ 'rice cake'
 4. tawtaw 'peanut'
 5. giati 'fish'
 6. baha 'meat'
 7. takutuy 'chicken'
 8. kuaykuay 'pig'
 9. malok 'dog'
 10. katin 'buffalo, cow', cf. Atayal katij
 11. gialim 'ghost, evil spirit'
 12. pilaʔ 'money' < PAN *pirak
 13. tamaku 'tobacco'
 14. ʔua 'none'
- ʔua pilaʔ 'to have no money'

12 Jia-rung Shie (謝佳榮) told me that there might be living Taokas speakers there, and helped to arrange for my field investigation.

15. mase 'many, much'
 pilaʔ mase 'to have much money'
16. ʔəman 'to eat' < PAN *k-um-aʔen
17. ka-mala 'to catch' < PAN *m-ala
 kamala takutuy 'to catch a chicken'
18. k'ane palala 'Don't talk'
19. muk'on gialim 'to see a ghost'
20. muk'on mahaliʔ kala 'to see pretty girls'
21. taokas ¹³ 'tribal name, self-appellation'
22. pulu 'Taiwanese, Fujien Chinese'
23. kisabaga 'Hakka'
24. doa 'two' ¹⁴ < PAN *DuSa
25. k'asab 'five', cf. Sai Lasəb, Paz xasəp, Babuza nahup, Hoanya hasip
26. gu-ituʔ 'seven' < PAN *pitu
27. makapal 'eight'
28. tanasuʔ 'nine'

Most of the above lexical items are the same or similar to the Taokas forms listed in Tsuchida (1982a).

The people there showed me a Taokas song book written in Chinese characters, but no one knew how to sing the songs or the meanings.

I was not at all surprised that I did not get much Taokas data in 1982. Half a century earlier, Miyamoto (1932) was only able to collect 46 lexical items based on the memory of a 77 year-old man and some other people. In the late sixties, Tsuchida (p.c.) was able to extract only one item 'wine' from a Taokas informant. It is very unlikely that anyone could get much more data for the Taokas language now. The best Taokas data available is

13 Actually the people there substituted -n for -s and produced the form as [taokan].

14 Like the Basay informant, the Taokas informants mixed up the order of the numerals and did not remember all of them.

about 400 items listed in Tsuchida (1982a), in which he discusses the close relationship among Taokas, Babuza, Papora, and Hoanya, based on shared phonological innovations. His conclusion is supported with evidence concerning the large number of shared cognates among these four languages in my own recent study (Li 1990).

3. The Survival of Formosan Languages

The following questions can be raised and discussed concerning the issues regarding the survival of Formosan languages.

(1) Can the extant Formosan languages still be saved from becoming extinct? How?

(2) What can be done to best document these languages?

(3) What priorities should be given to the study of these languages, considering the number of Formosan languages and limited man power or interested scholars?

A few Formosan languages are doomed to become extinct in a few years, including Pazeh, Thao, Kavalan, Kananavu and Saaroa. They are not actively spoken today. Only a few older people know these languages. To be more precise, only one old competent Pazeh speaker, less than ten Thao speakers, and a few dozen speakers for Kavalan, Kananavu and Saaroa are left. A few other languages, such as Saisiyat, will probably become extinct in a few decades. The above-mentioned Formosan peoples have adopted other languages: Pazeh, Thao and Kavalan peoples generally speak Taiwanese (a Chinese dialect), Kananavu and Saaroa peoples speak Bunun, and the Saisiyat people speak Atayal and Hakka (a Chinese dialect) in their daily life. Priorities, therefore, should be given to the study of the languages that will become extinct in a few years.

There are inadequate studies for all these languages. Most of the languages still lack very basic materials and data. Ideally each language should have at least the following materials: (1) a dictionary, (2) a reference grammar, and (3) texts. Native speakers can be trained in a short time to record their own language. Unfortunately the few speakers for these languages may be too old to learn writing skills. A reference grammar will take a lin-

guist two or three years. A dictionary will take even longer. However, combined efforts by a few interested scholars, graduate students and Formosan natives may produce some fruitful results.

Few Formosan languages have developed writing systems partly due to the discouragement of the Nationalist government of the Republic of China over the past four decades. As the country is becoming politically more liberal these days, some Formosan natives are seeking the means to preserve their cultural and linguistic identities. Various orthographic systems have been proposed for languages such as Amis and Paiwan. The Ministry of Education hastily set up a committee to study and devise orthographic systems for all the extant languages in September 1990. I was asked to head the project, with government officials, Formosan native missionaries and a few interested scholars involved. A draft of the orthographic systems for 15 Formosan languages was completed in May 1991 (Li 1991b). These are: Amis, Atayal, Sediq, Paiwan, Bunun, Puyuma, Tsou, Kanakanavu, Saaroa, Rukai, Saisiyat, Yami, Thao, Pazeh and Kavalan. Some peoples have become serious about the new writing systems for their languages, including Atayal, Sediq, Puyuma and Tsou. Amis and Paiwan are hesitant about our proposed writing systems as they have their own writing systems developed years ago (30 years ago for Amis). The main problem with the old Amis writing system is that it misses the important phonemic contrast between the pharyngealized (glottal) stop and the glottal stop. It is doubtful if any Pazeh, Thao, Kavalan, Kanakanavu and Saaroa people will start to write in their own native languages.

I hope that the data and findings presented here, limited though they may be, will encourage other scholars to investigate, record, and study these Formosan languages before they disappear.

(Accepted for publication 14 November 1991)

References

Abe, Akiyoshi

- 1930 Study of Aboriginal Languages. Taihoku. (In Japanese).
1938 Study of Place Names in Taiwan. Taihoku. (In Japanese).

Asai, Erin

- 1937 Investigation of languages in plain tribes. Nanpo Dozoku 4.3:55-56. (In Japanese).
1938 Investigations of the plain tribes by the Department of Linguistics, Taihoku Imperial University. Nanpo Dozoku 4.4:35. (In Japanese).
1939 Gravius's Formulary of Christianity in the Siraya language of Formosa, facsimile edition of the original of 1662. Taihoku Imperial University.

Campbell, Rev. William

- 1888 The Gospel of St. Matthew in Sinkang-Formosan, with corresponding version in Dutch and English edited from Gravius's edition of 1661. London: Trubner and Co.
1896 The Articles of Christian Instruction in Favorlang-Formosan. London: Trubner and Co.

Ferrell, Raleigh

- 1969 Taiwan Aboriginal Groups: Problems in Cultural and Linguistic Classification. Taipei: Institute of Ethnology, Academia Sinica monograph NO.17.
1970 The Pazeh-Kahabu language. Bulletin of the Dept. of Archaeology and Anthropology (BDAA), National Taiwan University 31/32: 73-97.
1971 Aboriginal peoples of the southwestern Taiwan plain. Bulletin of the Institute of Ethnology (BIE), Academia Sinica 32:217-235.
1982 Paiwan Dictionary. Pacific Linguistics C-73. Canberra.

Happart, Gilbertus

Paul Jen-kuei Li

- [1650] Word-book der Favorlangische Taal, ed. by W. R. Van Hoëvell.
Verhandelingen van het Bataviaasch Genootschap 18:31-381.

Li, Fang-kuei et al

- 1959 Notes on the Thao languages. BDAA 7:23-51. (In Chinese).

Li, Paul Jen-kuei

- 1976 Thao phonology. Bulletin of the Institute of History and Philology (BIHP),
Academia Sinica, 47.2:219-244.
- 1978 The case-marking systems of the four less known Formosan languages.
Proceedings of the Second International Conference on Austronesian
Linguistics, Fascicle 1, 569-615. Pacific Linguistics C-61.
- 1982 Kavalan phonology: Synchronic and diachronic. GAVA: Studies in
Austronesian Languages and Cultures Dedicated to Hans Kahler 17:479-495.
Berlin: Dietrich Reimer Verlag.
- 1990 Classification of Formosan languages: Lexical evidence. BIHP 61.4 : 813-
848.
- 1991a Early Austronesian terms for canoe and seafaring in the Formosan languages.
Paper presented at the Second International Symposium on Chinese
Languages and Linguistics, Taipei, August 9-11, 1991. Minzu Yuwen 74 :
14-17 (1992). (In Chinese).
- 1991b Orthographic Systems for Formosan Languages. Taipei : the Ministry of
Education. (In Chinese).

Li, Paul Jen-kuei and Lin, Ching-tsai

- 1990 Pazeh ritual songs for their ancestors and other types of songs. Field
Materials, Institute of Ethnology, Academia Sinica, Occasional Series 3:1-16.
(In Chinese).

Mabuchi, Toichi

- 1953 Migration and distribution of the Formosan aborigines(part 2). Minzokugaku
Kenkyu 18.4:23-72. (In Japanese).

Miyamoto, Nobuto

- 1932 Taokas vocabulary. Nanpo Dozoku 2.1:61-62. (In Japanese).

Moriguchi, Tsunekazu

- 1982 Some notes on accent in Kbalan : A minority aboriginal language in Formosa. Journal of Asian and African Studies 24 : 120-132.
- 1983 An inquiry into Kbalan phonology. Journal of Asian and African Studies 26 : 202-219.

Murakami, Naojiro

- 1933 Sinkan Manuscripts. Memoirs of the Faculty of Literature and Politics, Taihoku Imperial University.

Ogawa, Naoyoshi

- 1930 On Favorlang. Gengo to Bungaku 4:33-40. (In Japanese).
- 1939 Various aspects of the concept of time in Formosan languages. Minzokugaku Kenkyu 5.1:1-14. (In Japanese).
- 1943 On the words for "mortar" and "pestle" in Formosan languages. Minzokugaku Kenkyu, New Series 1.1:12-18, 1.2:38-46. (In Japanese).
- 1944 The position of Formosan languages in the Indonesian language group. Taiheiyoo Ken 1:448-494. (In Japanese).
- [n.d.] A comparative wordlist of Siraya dialects. Unpublished MS, printed in Tsuchida et al (1991).

Ogawa, Naoyoshi and Asai, Erin

- 1935 Myths and Traditions of the Formosan Native Tribes. Taihoku. (In Japanese).

Reid, Lawrence A.

- 1971 Philippine Minor Languages: Wordlists and Phonologies. Oceanic Linguistics Special Publication No.8.

Tsuchida, Shigeru

- 1969 Pazeh vocabulary. Unpublished ms.
- 1982a A Comparative Vocabulary of Austronesian Languages of Sinicized Ethnic

Paul Jen-kuei Li

Groups in Taiwan. Part 1: West Taiwan. Memoirs of the Faculty of Letters,
University of Tokyo No.7.

1982b Most persistent words in vanishing languages: The case of Papora (Taiwan).
GAVA' 17:463-477.

1985 Kulon: Yet another Austronesian language in Taiwan? Bulletin of the
Institute of Ethnology, Academia Sinica 60:1-59.

Tsuchida, Shigeru, Yukihiro Yamada and Tsunekazu Moriguchi

1991 Linguistic Materials of the Formosan Sinicized Populations I: Siraya and
Basai. University of Tokyo.

關於三種平埔族語言的新資料

(摘要)

李 壬 癸

台灣平埔族語言大多數在二十世紀初就已消失了，但今日偶爾仍然可以找到極少數老人依稀記得一些單字甚至片語。作者因一些熱心人士的指引，得以記錄台灣北部馬賽 (Basay) 族，南部西拉雅 (Siraya) 族，西部道卡斯 (Taokas) 族的一些語言資料。雖然並不多，但部分資料卻是前人沒有記錄過，而部分資料可以提供重要的線索，所以這些資料值得公諸於世。本文指出這些資料的價值和意義，並且列舉說明了馬賽語和西拉雅語的歷史音韻演變，以及它們和其他台灣南島語言的顯著差異。

最後一節全面檢討台灣南島語言面臨滅絕的危機，語言使用的現況，並提出挽救這些語言的具體方案——作者已為它們個別設計了文字書寫系統。如能鼓勵本族人儘量多講、多寫，相信至少可以延續一些語言的壽命，有助於保存這些語言和文化。