The Endangered Languages in Taiwan

Paul Li
Academia Sinica

Abstract

All Formosan languages are endangered and threatened with extinction. The problems of the criteria for being endangered languages and what aspects of language to study are briefly discussed. There are five most endangered Formosan languages: Pazih, Thao, Kanakanavu, Saaroa and Kavalan, each with a few speakers left. An account of the work that has been done on each of these languages is given. Finally there is a brief discussion of what has been done to conserve and revitalize all the endangered languages in Taiwan.

1. Languages in Taiwan: The Past and the Present

There are two main types of languages spoken in Taiwan, Austronesian and Chinese. All the aboriginal languages, spoken in the plains or mountain areas, belong to the Austronesian language family. The population of the indigenous peoples (450,000) is very small, only barely 2% of the total population of Taiwan (23,000,000). The rest of the population speaks a variety of Chinese dialects, including Mandarin, Minnan (Taiwanese), and Hakka. Mandarin has been adopted as the official language by the Nationalist Government since 1945, and it has become the predominant language at the expense of all the other languages in Taiwan. Hence these languages, especially the aboriginal, are endangered and threatened with extinction. Despite all the efforts of the new government in the past eight years, it seems hard to turn the tide.

When I started to work on Rukai, an Austronesian language in Taiwan, in 1970, it was sound and healthy, as were many other Formosan languages, such as Atayal, Seediq, Bunun, Tsou, Paiwan, and Amis. Even little kids spoke their native languages in the daily life. Unfortunately the situation is quite different today. All Formosan languages face the same problem of becoming extinct sooner or later. Few children and teenagers know much about their native languages; they all speak Mandarin. Even older people do not use their native language much in their own native village. There were plenty of competent native speakers with few researchers in the past, while there are far fewer native speakers with many more researchers today.
What is the main cause for the ill fate of these minority languages? First of all, the language policy and suppression of these languages by the Nationalist Government during the curfew period (1945-1987) did a lot of damage to them. Second, the mass media, especially TV, has made the linguistic situation deteriorate even more rapidly in the past few decades. It breaks one’s heart when one realizes the great value and importance of Formosan languages: (1) they are the most diverse in the entire Austronesian language family, and (2) they retain many archaic features, indispensable for the reconstruction of Proto-Austronesian.

The important issue of endangered languages in the world were recently discussed in Crystal (2000), Dalby (2002), and Robins and Uhlenbeck (1991). The main focus of this paper is documentation of the most endangered Formosan languages, the indigenous languages of Taiwan.

What can be considered an endangered language? We can set up certain criteria. One important criterion is whether it is being transmitted to the next generation, and if it will become extinct once all older speakers pass away. Unfortunately none of the Formosan languages is being transmitted to the younger generation. Another criterion is whether there are monolingual speakers. So far as I know, there are no more monolingual speakers in the 14 extant Formosan languages. Still another criterion is if there are only a small number of speakers of the language. But how small the number of speakers will make a language fall into the category of endangered languages?

All Formosan languages are threatened with extinction mainly because children do not learn to speak them or use them in their daily life. Four are most endangered as there are less than ten competent speakers left for each. These languages are: Pazih, Thao, Kanakanavu, and Saaroa. Next to these four languages is Kavalan, which still has a few dozen older speakers, all above middle age. In this paper I shall discuss fieldwork, documentation and publications of these languages in the past few decades.

What are the most important aspects of language we should study before it is gone? The vocabulary of a language may be limited. There are, however, an infinite number of sentences in any language. It is impossible to record “all” about a language. We do not really know what aspects of language are interesting and significant until we have a good understanding of it. The problem is that we simply do not have enough time to make a careful study of each highly endangered language.
Shigeru Tsuchida and I have been working on these languages, mostly independently, off and on over the past thirty to forty years. We have collected language data at all levels, vocabulary, phrases and sentences for grammatical analyses, texts, and even traditional songs. All texts and songs were tape-recorded, and a few Pazih, Thao and Kavalan texts were video-tape recorded. We give both interlinear glosses and free translation for all texts in computer files.

In the following section, I shall discuss the work that has been done on the most endangered Formosan languages, and suggest what needs to get done soon.

2. Documentation of the Most Endangered Formosan Languages

2.1 Pazih

Pazeh, formerly spoken in the central western plains of Taiwan, was first studied by Steere in 1873, Ino (1897), and Ogawa (1923) when it was still actively spoken, then by Ferrell (1970) and Tsuchida (1969), and still later by Li (1978, 1998) when it remained only in the memory of a few older people. Steere recorded a short wordlist and 28 sentences (see Steere 2002). Ino recorded a short text and a little vocabulary. Ogawa recorded a longer wordlist and gave a preliminary grammatical analysis based on the sentences he collected. Ferrell recorded some basic vocabulary, a little morphology and a short conversation. Tsuchida gathered the largest Pazih lexical data and two long texts from an 87 year-old man and two short texts from an old woman in 1969. Li recorded some lexical data and a few texts from various speakers. Li (1978) presented a preliminary analysis of the case-marking system in Pazih. Li (1998) is a more up-dated version for Pazih syntactic analysis, with some new findings, such as the Locative-focus and Referential-focus constructions, the locative forms of personal pronouns and aspect system. All of them, except Steere, Ino and Ogawa, had to work with the language in the memory of a few old people. Both Ferrell and Li have collected some data for the two closely related dialects, Pazih and Kaxabu.

We have come down to the very last speaker of Pazih at the age of 94 today. It is still one of the least studied Formosan languages. In addition to a sketch of grammar written in Chinese by Y. Lin (2000) and a paper on Pazih phonology and morphology by Blust (1999), two monographs were published, *Pazih Dictionary* (Li and Tsuchida 2001) \(^1\) and *Pazih Texts and Songs* (Li and Tsuchida 2001) \(^1\) and *Pazih Texts and Songs* (Li and Tsuchida 2001). It is reviewed by Blust (2003c) and Zeitoun (2001).
Words of traditional songs in Pazih were recorded by Li (Li and Tsuchida 2002:173-214), while musical notes were recorded and analyzed by Lin (Li and Lin 1990) and Wen (1998). Lori Su, a graduate student at National Tsing Hua University, has recently found that a noun stem can take the Patient-focus suffix –en to derive a verb in Pazih, as commonly found in Philippine languages. As it is a moribund language, it is hard to do any further significant investigation of it.

The introduction to *Pazih Dictionary* is a sketch of grammar, covering phonology (synchronic and diachronic), morphology (affixation and reduplication) and syntax (focus system, case markers, personal pronouns, aspect, imperatives, interrogatives, negatives, causative and nominalization), in 56 pages. Glosses for lexical forms and examples are given in both Chinese and English. It can serve as a reference grammar to Austronesian scholars.

### 2.2 Thao

The status of Thao has been ambiguous: It has been recognized neither as a mountain tribe nor as a plain tribe. Recent studies by Blust (1996) and Li (1998) indicate that Thao is more closely related with the four plains languages (Taokas, Babuza, Papora and Hoanya) formerly spoken in the western plains of Taiwan.

The first important linguistic study of the Thao language was carried out by Fang Kuei Li (1956), “Notes on Thao Language,” which gives a grammatical sketch and a useful wordlist, and has become the only source for citation by comparative Austronesian scholars, such as Isidore Dyen, Shigeru Tsuchida and Robert Blust, for many years. Twenty years later Li (1976) published his “Thao Phonology,” which adopts a generative approach and emends F. K. Li’s (1956) sound system. Another two decades elapsed before Blust's (1996, 1998) papers on Thao appeared. Blust (1998) accounts for some of the phonological problems Li did not deal with or did not deal with in a satisfactory manner. Blust (1996) refutes Li’s (1990) grouping Thao with Bunun, as based on lexical evidence, and proposes to group Thao immediately with the western plains languages, Taokas and Babuza. Li’s (2001) own recent study indicates that Blust is justified in refuting his grouping Thao with Bunun, but not well justified in suggesting that Thao is more closely related to Taokas and Babuza than any other Formosan languages, including Papora and Hoanya in the western plains of Taiwan. Tsuchida has also done some fieldwork on Thao and given a sketch of Thao grammar in his (Tsuchida 1989) short article.
Thao fares a little better than Pazih. There are still about ten speakers of Thao, mostly above the age of 70. In addition to the two most important works on Thao, *Thao Dictionary* by Blust (2003) and a PhD dissertation *An Ergative View of Thao Syntax* by Shan-shan Wang (2004), there are a few MA theses and papers on the Thao language, including Laura Chang’s (1998) paper *Thao reduplication*, Weng’s (2000) *A Contrastive Study of Tense, Mood and Aspect Systems in Tsou and Thao*, Y. Chen’s (2000) *Negation in Thao and Tsou* and Lu’s (2003) *An Optimality Theory Approach to Reduplication in Formosan Languages*. Blust’s *Thao Dictionary* is extremely copious and rich in data. Wang’s dissertation is by far the most comprehensive syntactic study of Thao to date, covering its basic clause structure, topicalized constructions, deictics, and structure of noun phrases, and so on. Blust (1998, 2001) has a squib on the Thao patient focus perfective and another squib on Thao triplication. I have collected dozens of texts, with interlinear glosses and free translation over the past three decades, but still unpublished.

### 2.3 Kanakanavu

Among all extant Formosan languages, Kanakanavu and Saaroa are the least studied and most poorly understood. There is no reference grammar, no dictionary, no thesis or dissertation on either of the languages. These two languages are closely related, but their linguistic position is still unsettled; see, for example, Chang (2006), who argues that morphosyntactic evidence indicates that Kanakanavu and Saaroa may not subgroup with Tsou.

Except for the brief descriptions and data given in Ogawa and Asai (1935:721-739) and Tsuchida (1976:26-58), not much has been done or published on Kanakanavu. Mei’s (1982) paper “Pronouns and verb inflection in Kanakanavu” is a welcome contribution to the field. Ho (1997) published a sketch of Kanakanavu grammar, while Li (1997) published one on Saaroa, both with descriptions of their phonology, morphology and syntax, in Li et al’s (1997) monograph. Szakos has done extensive fieldwork and collected a large amount of data for both Kanakanavu and Saaroa, but has not published anything on the languages as yet. There are not many competent speakers left, so it is urgent to investigate these two languages more carefully and in-depth before they disappear.

---

2. It is reviewed by Zeitoun (2005).

3. Thao reduplication has also been studied by Lu (2003) and Lee (2007).
There are only 7 competent speakers of Kanakanavu today. The few publications on the language are a sketch of grammar and 7 texts in Ogawa and Asai (1935:723-739), a description of its grammatical structure in Tsuchida (1976:26-58), a sketch of grammar in Chinese by Ho (Li et al 1997:228-63), Mei’s (1982) paper, and the monograph *Kanakanavu Texts* by Tsuchida (2003). Mei, Li, and Szakos have each collected some texts, but none are published as yet. Folk songs in both Kanakanavu and Saaroa were published in CDs with a pamphlet (Wu et al 2001), a successful collaboration of a linguist and an ethnomusicologist.

### 2.4 Saaroa

Similar to Kanakanavu, there are about 10 competent speakers of Saaroa. The few publications on the language are: a sketch of grammar and 8 texts in Ogawa and Asai (1935:695-719), a description of its grammatical structure in Tsuchida (1976:59-83), a sketch of grammar in Chinese by Li (Li et al 1997:272-94), a study of its traditional folk songs by Li (2006), and two papers on syntax or semantics by Radetzky (2004, 2006). Radetzky has investigated the language in the past few years with some interesting observations.

It is obvious that a lot more serious linguistic work needs to be done on both Kanakanavu and Saaroa as soon as possible. Two graduate students of National Tsing Hua University, Chun-ming Wu has been doing some fieldwork on Kanakanavu, while Chao-lin Li has been doing some fieldwork on Saaroa.

### 2.5 Kavalan


Li’s (1996:55-162) monograph contains a long chapter which is perhaps the most comprehensive study of the Kavalan language. It deals with phonology,
morphology, syntax and gives comparative word lists for Kavalan dialects. Moreover, it provides not only 12 texts collected by the author himself, but also 11 texts collected for various Kavalan dialects by Asai in 1936. Liao’s (2002) paper “The interpretation of tu and Kavalan ergativity” and her (Liao 2004) PhD dissertation, *Transitivity and Ergativity in Formosan and Philippine Languages*, which includes Kavalan as one of four languages she treated, are based on the texts collected by Li. Liao’s dissertation is a thorough discussion of transitivity and ergativity in these languages, including Kavalan.

*Kavalan Dictionary*, coauthored by Li and Tsuchida (2006), contains all lexical items and many examples collected by the authors, Ogawa, their colleagues and students. The introduction is a reference grammar, which covers phonology, morphology, and syntax. The following two new observations are made: One is that there are geminate consonants, and the other is the noun-verb distinction, which is probably due to an earlier stress pattern: Nouns would have had stress on the ultima and verbs on the initial syllable of lexically related pairs, e.g. *btu* ‘stone’ vs. *battu* ‘to throw with a stone’.

Y. L. Chang’s (1997) Ph.D. dissertation is a “government and binding” approach to two Formosan languages, Kavalan and Seediq. It is theoretically oriented and provides many interesting and insightful observations on the languages under study, as does the paper "Actor-sensitivity and obligatory control in Kavalan," co-authored by Y. L. Chang and Tsai (1998). His (Chang 2006) paper “The guest playing host: Adverbial modifiers as matrix verbs in Kavalan” discusses the general property of Formosan languages, in which adverbial modifiers function as main verbs syntactically.

It is clear that Kavalan is one of the best studied among these five most endangered Formosan languages.

3. Conservation and Revitalization of the Endangered Languages

What can be done to conserve and revitalize the endangered languages? Most natives do not seem to care. A few who are concerned do not know what they can do because younger people show little interest in learning their own native languages. There is little incentive for learning them. Despite the encouragement and millions of dollars spent by the government, there are few good language textbooks for these languages, as most language textbooks were prepared by non-professionals. Moreover, there are few well trained language teachers. Children and younger people would soon lose interest even if they tried.
We encourage all Formosan natives to use their own native languages at home and in their own speech communities as much as possible, but to no avail. The language policy of the current government requires that native language courses be offered and that pupils in elementary and junior high school take at least one of the three native languages, Minnan, Hakkha, or a Formosan language.

The Council of the Indigenous Affairs of the Taiwan government has been giving indigenous language proficiency tests to encourage natives to learn their own languages in the past decade or so. Thousands of people have taken the tests and got certificates over these years. Unfortunately there was not enough incentive for more people to do so. A good incentive was most recently provided: Any native who passes a language proficiency test will have the privilege of earning an extra 35% increase to the total marks (s)he gets when (s)he takes an entrance examination to senior high school or college. More than ten thousand pupils or students were registered for the indigenous language proficiency test and about an average of 75% from each ethnic group passed it in early 2007.

Orthographic systems were devised for all the fourteen extant Formosan languages and Yami, and they were officially announced by the Ministry of Education in 1991 (Li 1991). Slightly modified systems for the orthography of some of the languages were officially adopted jointly by the Ministry of Education and the Council of Indigenous Affairs in December 2005.

Mandarin Chinese is the only officially recognized national language in Taiwan. Hence it is the only dominant language, and has steadily been encroaching on the sphere of other languages, including the Chinese languages (or dialects) spoken in Taiwan, Minnan, Hakkha, and the indigenous languages. Ever since the Democratic Progressive Party came into power in 2000, the language policy of the government has been to encourage pupils in elementary school and students in junior high school to learn at least a native language. Despite all the efforts of the new government and the people involved, it is hard to turn the tide. It is most likely that one language will die after another and all Formosan languages will become extinct in a few decades.
References

Blust, Robert


Chang, Laura M. (張美智)


Chang, Henry Yung-li (張永利)


Chang, Yung-li & Wei-tien Dylan Tsai (張永利，蔡維天)


Chen, Youmehim (陳佑民)


Crystal, David


Dalby, Andrew

Ferrell, Raleigh

Ho, Dah-an (何大安)
1997 〈卡那卡那富語〉[The Kanakanavu language], 李壬癸編著《高雄縣南島語言》[Formosan language in Kaohsiung], 228-271。

Huang, Lillian

Inô Kanori（伊能嘉矩）
1897 番語集[Pazeh language]. Field notes.

Lee, Amy Pei-jung（李佩容）

Li, Fang-kuei, Chi-lu Chen and Mei-chün Tang（李方桂，陳奇祿，唐美君）
1956 〈邵語記略〉[Notes on the Thao language], 《國立台灣大學考古人類學刊》BDAA 7:23-51。

Li, Paul Jen-kuei（李壬癸）
1983 Notes on Thao dialects. 《考古人類學刊》BDAA 43:48-50.
1991 *Orthographic Systems for Formosan Languages.*
1996 《宜蘭縣南島民族與語言》[The Formosan Tribes and Languages in I-Lan]。宜蘭[I-Lan]: 宜蘭縣政府[The Government of I-Lan]。
1997 《高雄縣南島語言》[The Formosan Languages in Kaohsiung]。高雄[Kaohsiung]: 高雄縣政府[The Government of Kaohsiung]。
1998 《巴則海語的格位標記系統》[The case-marking system of the Pazeh language], 《台灣語言及其教學國際研討會論文集》[Papers from the
International Symposium on the Languages and Language Teaching in Taiwan] 1:57-81。新竹師院[Hsinchu Normal College]。


2001〈邵族的地位—兼評白樂思（Blust 1996）的邵族地位說〉[On the linguistic position of Thao—Some remarks on Blust's (1996) ‘Some remarks on the linguistic position of Thao’], 詹素娟 (Su-chuan Jan)、潘英海 (Ying-hai Pan)编，《平埔族群與台灣歷史文化論文集》[Papers Presented at the Symposium on the Plains Aborigines and Taiwan History], 165-184。中央研究院台灣史研究所籌備處[Institute of Taiwan History Preparatory Office, Academia Sinica]。

2006〈沙阿魯阿的歌謠〉[Saaroa Songs]，《山高水長：丁邦新先生七秩壽慶論文集》(Mountain Lofty, River Long: Festschrift in Honor of Professor Pang-hsin Ting on His Seventieth Birthday)。

Li, Paul J. K. and Ching-tsai Lin (李壬癸，林清財)

1990〈巴則海族的祭祖歌曲及其他歌謠〉[The Pazeh ritual songs for ancestors and other types of folk songs], 《民族學研究所資料彙編》FMIE 3:1-16。

Li, Paul J. K. and Shigeru Tsuchida (李壬癸，土田滋)


Liao, Hsiu-chuan（廖秀娟）

2002 The interpretation of tu and Kavalan ergativity. OL 41.1:140-158.


Lin, Ju-en（林主恩）


Lin, Ying-chin（林英津）

2000 《巴則海語》[Pazih language]。台北[Taipei]：遠流出版公司[Yuan–liu Publishing Co.]。

Lu, Michael Shun-chieh（呂順結）

2003 An Optimality Theory Approach to Reduplication in Formosan Languages.

Mei, Kuang (梅廣)


Moriguchi, Tsunekazu (森口恒一)


Ogawa, Naoyoshi (小川尚義)

1923 〈台灣の蕃語に就て〉[On the Formosan aboriginal languages],《台灣時報》[Taiwan Jihô] 49:6-23 (1923:10)。漢譯文見黃秀敏譯、李壬癸編審（1993）。

Ogawa, Naoyoshi and Asai, Erin (小川尚義，淺井惠倫)

1935 《原語による台灣高砂族傳說集》[The Myths and Traditions of the Formosan Native Tribes]。台北：台北帝國大學言語學研究室 [Taihoku: Taihoku Teikoku Daigaku Gengo-gaku Kenkyu-shitsu]。

Radetzky, Paula Kadose


Robins, R. H. and E. M. Uhlenbeck, eds.


Shen, Chia-chi (沈嘉琪)


Steere, Joseph Beal

2002 Formosa and Its Inhabitants, edited by Paul Li. Institute of Taiwan History (Preparatory Office), Academia Sinica.

Tsai, Dylan Wei-tien (蔡維天)

1997 〈台灣南島語疑問詞的無定用法—噶瑪蘭語、鄒語及賽德克語的比較研究〉[Indefinite Wh-construals in Formosan languages—A contrastive study in Kavalan, Tsou and Seediq],《清華學報》[Tsing Hua Journal of Chinese Studies] 新 27.4:381-422。

Tsuchida, Shigeru (土田滋)

1976 Reconstruction of Proto-Tsouic Phonology. Tokyo: Study of Languages &
Cultures of Asia & Africa, Monograph Series No.5, Tokyo University of Foreign Studies.

1989 〈サオ語〉[The Thao language]，《言語學大辭典》[Encyclopaedia of Linguistics] 2:15-16。東京[Tokyo]：三省堂[Sanseido]。

1993 〈カバラン語〉[The Kavalan language]，《言語學大辭典》[Encyclopaedia of Linguistics] 5:89-99。東京[Tokyo]：三省堂[Sanseido]。


Wang, Shan-shan（王珊珊）


Wen, Chyou-chu（溫秋菊）

1998 〈巴則海族祭祖歌“ai-yan”初探〉[A preliminary study of Pazeh's 'ai-yan: The songs for the rite in honor of ancestors]，《藝術評論》[Arts Review] 9:45-84。

Weng, Tsui-hsia（翁翠霞）

2000 〈鄒語和邵語的時、態、貌系統之比較與研究〉[A contrative study of tense, mood and aspect systems in Tsou and Thao]，嘉義：中正大學語言所碩士論文 [Unpublished MA thesis, Chungcheng University]。

Wu, Rung-shun et al（吳榮順等）

2001 〈南部鄒族民歌〉[Songs of Hla'alua and Kanakanavu]。台灣原住民音樂紀實 9。台北[Taipei]：風潮[Wind Records]。

Yeh, Yu-Ting（葉俞廷）


Zeitoun, Elizabeth 齊莉莎
