

Retrospect and Prospect of Recent Studies on Extinct Pepo Languages

TSUCHIDA Shigeru

Background

About 15 years ago, I gave an informal speech on Pepo studies at the Center of Taiwan History Field Research, which was later published as Tsuchida (1992).

Since then, a large number of remarkable studies on Pepo languages have been achieved by various scholars and researchers. Today I will report briefly about the significant progress in recent years. Thao, Kavalan, and Pazeh used to be classified under Pepo. We usually understand by the name Pepo that they are heavily and severely sinicized so much that their languages are not spoken any more. Actually, however, at least Thao and Kavalan are still actively spoken. Therefore I will treat only Pazeh here.¹

1. Recent Findings of New Data for Pepo Languages

1.1 Asai and Ogawa Materials at AA-Ken and in Nanzan University

Ogawa Naoyoshi is the founder of the research studies of, not only the Taiwanese and Hakka languages, but also the Formosan Austronesian languages, and for the second field, Asai Erin is the next to him. They left many notebooks and manuscripts on these languages, but nobody knew about them until Asai passed away in 1969, and subsequently Mrs. Erin Asai sold all of them to the Research Institute of Asian and African Languages and Cultures, the Tokyo University of Foreign Studies (hereafter abbreviated as AA-Ken).

When I was at AA-Ken, I quickly checked the materials and utilized some of them in my own studies, but I didn't have enough time to put them into order or make a list of the

¹ Professor Li Jen-kuei read an earlier version of this paper, and gave me very valuable suggestions for improving it.

materials. Later I moved to Tokyo University, and had no chance to use them since then. Toward the end of the last century, it was decided that AA-Ken would be moved to the present location from the old campus, and every member of the Institute was busy throwing away stuff to the corridors. Among them were found some very old-looking notebooks and sheets of paper piled up in the corner of old shabby bookcases. Someone took a last look at the garbage to be thrown away, and thought they might be invaluable old materials on some Formosan languages.

When that news came to me, I was much surprised that they were treated in such a terrible situation. After I left the AA-Ken, there was nobody specialized in the Formosan Austronesian languages, and nobody was interested in the old dusty materials. At that time AA-Ken also needed some eye-catcher articles to attract people's attention to show the importance of the Institute. And thus started a project of putting in order all the materials left and classifying them to make a list. Almost at the same time I heard the news that there were found several carton boxes of materials left by the late Professor Asai at the Institute of Anthropology, Nanzan University in Nagoya, to which Professor Asai was affiliated before he died (see Li 2000a). Professor Li Jen-kuei spent lots of time preparing lists of the materials at these two institutions.

The materials left by Asai and Ogawa are of four types: (1) notebooks and sheets of paper of the Formosan Austronesian languages, (2) photos, (3) records, and (4) land deeds of Pepo usually known as 新港文書. The lists of the notebooks and records appear in Mio and Toyoshima (2005), and that of Sinkang manuscripts is in Li (2002a:192).

1.2 New Findings of INO Kanori's Notebooks on Pepo Languages

In 1995 Prof. KASAHARA Masaharu got acquainted with Mr. EDA Akihiko, one of the descendants of the INO family, and was informed that all the materials left by INO Kanori had been deposited in the City Museum of Toono (遠野), Iwate Prefecture (岩手県). Toono is INO's home town. Some anthropologists, linguists, and historians and other interested people visited the Museum in June 1996, and found many valuable materials. All the linguistic notebooks written by INO were put in order by MORIGUCHI Tsunekazu, and were published in Moriguchi (1998). Most of the data were already compiled by OGAWA in his own notebooks, and those in the western plains, i.e. Taokas, Papora, Babuza, and Hoanya

were already made available in Tsuchida (1982).

1.3 Findings of New Sinkang Manuscripts

After Murakami's monumental transcriptions and editing in 1933 of all the Sinkang manuscripts known to be available by that time, quite a number of new Sinkang manuscripts have been reported. The most significant discovery of all is those manuscripts collected by Joseph Beal Steere, professor of zoology, the University of Michigan, who spent six months from October 1873 to March 1874 in Taiwan, and happened to find and purchased 29 manuscripts during his trip. He was the first person who reported the existence of such handwritten manuscripts. Those documents were later called Sinkang manuscripts. Unfortunately most of them were unknown until Prof. Henry Wright of the University of Michigan sent to Prof. Li Jen-kuei "some 20 photocopies of the Sinkang manuscripts to determine whether the originals were of any value and worth conserving."(Li 2002b) With a great effort, Prof. Li made two trips to Ann Arbor to publish Steere's monograph entitled "Formosa and Its Inhabitants" (Li 2002a) as well as the 16 new manuscripts found in Michigan (Li 2002b).

In Taiwan, on the other hand, there have been found a number of new Sinkang manuscripts. Besides those 3 reproduced in Tseng and Tung (1999) and dozens in Chen & Tsai (2006), I have heard personally from a couple of people that more and more new Sinkang manuscripts have been procured by some wealthy dilettanti in recent years. I sincerely hope that they would not keep them under lock but open them to serious researchers of Siraya and other closely related languages in the near future.

2. Analyses of the Data and the Results of the Studies

The most remarkable progress in the field of the studies of Pepo languages is perhaps the problem of their mutual relationships and subclassification.

2.1 Subclassification of the Western Plains Pepo Languages

Tsuchida (1982) arranged the Ogawa's data on Taokas, Papora, Babuza, and Hoanya, and based on the phonological renovations and the commonly shared lexical features, I

realized the clustering of Taokas and Papora on one hand while Babuza and Hoanya on the other. Blust (1996) added Thao to this group (Western Plains). Li Jen-kuei (2001b, 2003b) further suggests that Pazeh is more distantly related to Thao and other four Western Plains group and named this large group Northwestern.

When I was arranging the linguistic data of Ketangalan and Basai left by Ino, Ogawa, Asai, and some other people, I found that the words collected at the former villages of Ku-lon (龜崙) in Tao-yuan-hsien 桃園縣) are quite different from the surrounding Ketangalan language, but rather show the same phonemic change as that in Saisiyat (Tsuchida 1985). Of course we cannot say anything certain based only upon a comparison of a few lexical items, but unfortunately there is only very little linguistic data on Katangalan and Kulon, and we cannot expect any more new data would become available, we will have to be satisfied with saying the Kulon is more closely related to Saisiyat rather than Ketangalan.

The most active scholar in the studies of Ketangalan and Basai is Prof. Li Jen-kuei. He utilized linguistic data left by Ogawa and other people, and proposes to use the name Ketangalan as the cover term including Luilang, Basai, Torbuan, and some others. Thus according to Li (1995), Ketangalan is divided into two large groups: Southwestern (Luilang, Nankang, Puting)² and Northeastern (Basai, Torbuan, Linao). It is very regrettable that we have extremely limited data for Qauqaut, so that we cannot say anything clear about the position of this language. Prof. Li, on the other hand, proposes a close relationship between Basai and Kavalan, on one hand, and Amis and Siraya on the other, and these four languages belong to East Formosan group (Li 2001a, 2004), as originally suggested by Blust (1999b).

In southern Taiwan three Pepo languages were spoken: Siraya, Taivuan, and Makatau. Tsuchida et al. (1991) arranged the data of these three languages left by Ogawa and others, and compiled a comparative vocabulary. Professor Li Jen-kuei called them Sirayaic as a cover term, and based on the phonological and lexical innovations he concluded that Siraya was the first split from the Sirayaic group, while Taivuan and Makatau are more closely related (Li (in prep.)).

	PAn	Taivuan	Siraya	Makatau
(1)	*l	zero ~ h	r	r

² According to my own personal studies, it seems that putting contains a number of Kavalan lexical elements. I do not know the reason. It is worth studying more carefully with this language.

(2)	*N	l	l	n
(3)	*D,*d	r ~ d	s	r ~ d
(4)	*k	zero	-k-	-k-
	*-S,*R	zero	-g-	-----
(5)	(future)	-ah	-ah,-ili	-ani

The most conspicuous and at least to me the most interesting finding by Professor Li is that Matau (麻豆) was, unlike both Ogawa and Tsuchida wrongly thought, not Siraya but Taivuan. And also it is interesting to note that the Gospel of St. Mathew and the Formulary were written in some dialect of Taivuan, whereas the Utrecht manuscripts reproduced first in Van der Vlis (1842) and quoted in full in Murakami (1933) were written in Siraya. Thus:

Pan	St.Matthew	Formulary	Utrecht Ms.	English
*D	r,d	r,d	s	
*DapaN	rahpal	rahpal	sapal	`foot'
*DuSa	dou-rouha	rou-rouha	so-soa	`two'
*likuD	rikour	rikour	ricos	`back'
*-k-	-k-, zero	zero	-k-, zero	
*(i)aku	jau, -au	jau	iau, -au	`I'
	akousi, aousi	aoussi	acoussey	`not have'
*S, *R	-h-, zero	-h-, zero	-g-, zero	
*kaSu	kow	kow	cau	`thou'
*DuSa	dou-rouha	rou-rouha	so-soa	`two'

2.2 Progress of Studies in Siraya and Sirayaic languages

A great deal of progress has been done in Siraya studies. Some important studies on Siraya grammar are a series of Adelaar's papers (Adelaar 1997, 1999, 2000, 2004, 2006), and Tsuchida's (Tsuchida 1996, 2000).

The transcription of Sinkang manuscripts is another important contribution to Siraya studies. Besides Professor Li's readings of the newly found Sinkang manuscripts (Li 2002b), Professor Weng Jia-yin paid very much effort to transcribing several Sinkang manuscripts (Weng 1989a, 1989b, 1990a, 1990b). And Wu & Weng (2006) should not be missed here as well. English translation of the Dutch missionary works by Chen Bien-horn (陳炳宏 2001,

2005) is another great contribution to our Siraya studies. A complete dictionary based on these digitalized data should be compiled, and I understand Professor Li has been working on this project. We expect it to become available hopefully in the near future.

2.3 Pazih

The Pazih language is on the very verge of extinction with only one competent speaker of 93 years old as of 2007. However, I believe that Pazih was lucky enough to have at least one dictionary (Li & Tsuchida 2001), a collection of texts and traditional songs (Li & Tsuchida 2002), a grammar book (Lin 2002), and several papers of good quality on some aspects of its grammar (Blust 1999a; Li 2000, 2002). All other extinct Pepo languages never could have this much luck.

2.4 Favorlang Vocabulary

Favorlang is actually almost the same as Babuza, or we should rather say Babuza is the descendant of Favorlang. Side by side with Siraya, this language was fortunately relatively well documented by the Dutch missionaries, but somehow serious analyses of this language have never been tried. Professor Li Jen-kuei edited the data left by Ogawa, and published English-Favorlang Vocabulary (Li 2003a) with his introduction, 'Notes on Favorlang, an extinct Formosan language.' Favorlang-English dictionary is already available by Medhurst (1840), so that these two books will be of great help for future studies of this less known language. Only even simply to digitalize the dictionary of Medhurst would be much helpful for us to analyze the language.

2.5 Basay

In spite of very little amount of the data available of Basay, Professor Li Jen-kuei fully utilized the fieldnote scratches left by E. Asai, and tried to pick up and analyze several important linguistic aspects such as the pronouns, case markers, focus system, aspect system, negation, and interrogatives (Li 1996, 1999). Based on these data, as well as the phonological and lexical comparisons, he concluded that Basay is a Formosan language, closely related to Kavalan and also to Amis, with a lexical influence of some Philippine

language(s) perhaps through those Philipino sailors hired by the Spaniards.

2.6 Other Pepo Languages

Professor Li Jen-kuei edited a large amount of data left by Ogawa and Asai, and published *A Comparative Vocabulary of Formosan Languages and Dialects* by Naoyoshi Ogawa (Li 2006). This huge book includes 286 lexical entries from all the languages and many dialects of the Formosan Austronesian languages known until early 1920's. It seems to me that this comparative vocabulary was compiled by Ogawa in order to understand and roughly grasp the relationships of the Formosan Austronesian languages before Ogawa himself started his own fieldwork. Therefore it contains not only Pepo languages but also all other Formosan Austronesian languages, and most of them have been outdated by now. But the data of Pepo may be still of some use. What is more, some sounds which disappeared in the present language(s) and/or dialect(s) can be inferred, based on the letters and characters left in old documents collected in the earlier dates (see Tsuchida forthcoming; cf. also Tsuchida (2008)).

3. Problems Left

- 1) As mentioned above, we need to study Favorlang more carefully, including the texts and morphology of the vocabulary left by the Dutch missionaries. As of mid-May of 2008, Professor Li Jen-kuei's project for two years (Aug.2008--Jul.2010) was fortunately decided to be funded by CCKF. We can expect to work out both Favorlang and Basay languages.
- 2) Ino first reported on Arikun, Lloa, and Hoanya, but later Ogawa (Ogawa and Asai 1935) treated those three languages under the cover term Hoanya because, he says, the lexical data left are rather similar even though he admits the paucity of the data. Tsuchida (1992) presented the problem why they had different self-appellations if they spoke the same language. If we treat the lexical data more carefully, we might be able to screen out the differences between these three languages or dialects.

References

- Adelaar, K. Alexander. 1997. 'Grammar notes on Siraya, an extinct Formosan language.' *Oceanic Linguistics* 36(2): 362-397.
- . 1999. 'Retrieving Siraya phonology: a new spelling for a dead language.' E. Zeitoun and P. J.-K.Li (eds.) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*. pp. 313-354. Symposium Series of the Institute of Linguistics, Academia Sinica, No.1
- . 2000. 'Siraya reduplication.' *Oceanic Linguistics* 39(1): 33-52.
- . 2004. 'The coming and going of 'Lexical Prefixes' in Siraya.' *Language and Linguistics* 5(2):333-361.
- . 2006. 'Siraya dialogue.' Chang et al. 2006.10: 665-685.
- Blust, Robert. 1996. 'Some remarks on the linguistic position of Thao.' *Oceanic Linguistics* 35: 272-294.
- . 1999a. 'Notes on Pazeh phonology and morphology.' *Oceanic Linguistics* 38(2): 321-365.
- . 1999b. 'Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics.' In E.Zeitoun and P.Li (eds.), *Selected Papers from the Eighth International Conference on Austronesian Linguistics* 37.2:277-297.
- Chen, Bien-horn 陳炳宏. 2001. 〈台南 SIRAYA 語馬太福音字彙解讀〉 (Notes on Vocabulary of Siraya of Matthew). xvii+416pp., 台北：台灣平埔原住民協會。
- . 2005. 〈臺(南)SIRAYA 語基督教信仰要項解讀〉 (Notes on Formulary of Christianity in Formosan Siraya Dialect with translation in English, Chinese and Dutch H. 838pp., Private Circulation.
- Chen Chiu-kun and Tsai Cheng-wei (eds.) 陳秋坤・蔡承維 (編著). 2006. 《大崗山地區古契約文書匯編》。中央研究院臺灣史研究所・高雄縣政府文化局。
- Li, Paul Jen-kuei. 1995. 〈台灣北部平埔族的種類及其互動關係〉 [Sinicized tribes in northern Taiwan and their interactions] 潘英海・詹素娟 編 [Pan, Ying-hai and Chan, Su-chuan, eds.], 《平埔研究論文集》 [Symposium of the P'ingpu Studies], 21-40. 台北 [Taipei]: 中研院台史所籌備處 [Institute of Taiwan History Preparatory Office, Academia Sinica]。
- . 1996. 《宜蘭縣南島民族與語言》。宜蘭：宜蘭縣政府。

- . 1999. 'Some problems in the Basay language.' In E. Zeitoun and P. Li (eds.), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, pp. 635-664. Symposium Series of the Institute of Linguistics (Preparatory Office), No. 1. Taipei: Academia Sinica.
- . 2000a. 'Formosan language materials by Ogawa at Nanzan University.' 《臺灣史研究》 5.2: 147-158. Reprinted in 『南山大學人類學研究所通信』 8:2-7 as 「南山大學所藏小川尚義による臺灣原住民諸語資料」.
- . 2000b. 'Some aspects of Pazez syntax.' In *Grammatical Analysis: Morphology, Syntax, and Semantics. Studies in Honor of Stanley Starosta*. (eds. V. P. De Guzman and Byron W. Bender). pp. 89-108. Honolulu: University of Hawai'i Press.
- . 2001a. 〈巴賽語的地位〉 [The linguistic position of Basay], 《語言暨語言學》 [Language and Linguistics] 2.2: 155-171.
- . 2001b. 〈郡族的地位—兼評白樂思(Blust 1996)的郡族地位說〉, (The linguistic position of Thao with some remarks on Blust's (1996) paper on Thao), Zhan Su-juan and Pan, Ying-hai (eds.) 詹素娟·潘英海 (編) 《平埔族群與台灣歷史文化論文集》, 台北: 中央研究院台灣史研究所籌備處。 pp. 165-184。
- (ed.). 2002a. Joseph Beal Steere. *Formosa and Its Inhabitants*. xxi+217pp. 台北: 中央研究院臺灣史研究所籌備處。
- . 2002b. 〈新發現十五件新港文書的初步解讀〉 [Preliminary interpretations of the 15 recently uncovered Sinkang manuscripts] 《臺灣史研究》 [Taiwan Historical Research] 9.2: 1-68.
- . 2002c. 'Nominalization in Pazez.' *Language and Linguistics* 3.2: 227-240.
- (ed.). 2003a. *English-Favorlang Vocabulary by Naoyoshi Ogawa*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Asia-African Lexicon Series No. 43. 224pp.
- . 2003b. 'The internal relationships of six western plains languages.' *Bulletin of the Department of Anthropology* 61: 39-51.
- . 2004. 'Origins of the East Formosans: Basay, Kavalan, Amis, and Siraya.' *Language and Linguistics* 5.2: 363-376.
- . 2006.08. 〈台南和高屏地區的平埔族語言: 兼論麻豆社的地位〉。《建構西拉雅研討會論文集》 (葉春榮主編) pp. 17-38。台南: 台南縣政府。
- . (in prep.) 'Linguistic differences of Siraya, Taivuan and Makatau.' In Adelaar and

Pawley (eds.)

- Li, Paul Jen-kuei, and Masayuki Toyoshima 李壬癸・豊島正之 (eds.). 2006.03. 〈臺灣蕃語蒐錄 by 小川尚義〉 (A Comparative Vocabulary of Formosan Languages and Dialects by Naoyoshi OGAWA), Asian and African Lexicon Series No.49. xxiii+714pp. 東京:東京外国語大学アジア・アフリカ言語文化研究所.
- Li, Paul Jen-kuei, and Shigeru Tsuchida. 2001. Pazih Dictionary. Language and Linguistics Monograph Series Number A2. xii+391pp. Taipei: Institute of Linguistics, Academia Sinica.
- . 2002. Pazih Texts and Songs. Language and Linguistics Monograph Series Number A2-2. 258pp. Taipei: Institute of Linguistics, Academia Sinica.
- Medhurst, W.H. 1840. Dictionary of the Favorlang Dialect of the Formosan Languages, by Gilbertus Happart, written in 1650. 383pp.
- Mio, Yuko, and Masayuki Toyoshima (eds.) 三尾裕子・豊島正之 (編). 2005. 『小川尚義・浅井恵倫 台湾資料研究--浅井・小川未整理資料の分類・整理・研究プロジェクト』 東京:東京外国語大学アジア・アフリカ言語文化研究所. 395pp.
- Moriguchi, Tsunekazu (ed.) 森口恒一 (編).1998.《伊能嘉矩蕃語調査ノート》(INO Kanori, Field Notes of Formosan Aboriginal Languages)。台北：南天書局。[iv+]228+9+vii pp.
- Murakami, Naojiro 村上直次郎. 1933. Sinkan Manuscripts (新港文書). Memoirs of the Faculty of Literature and Politics, Taihoku Imperial University Vol.2, No.1. (台北帝国大学文政学部紀要第二卷第一号) 台北.
- Ogawa, Naoyoshi, and E.Asai 小川尚義・浅井恵倫.1935. 《原語による臺灣高砂族傳説集》(Myths and Traditions of the Formosan Native Tribes)。台北：台北帝国大学言語学研究室.
- Tseng, Cheng-ming, and Yuan-Chao Tung (ed.) 曾振名・童元昭. 1999. 《噶瑪蘭・西拉雅古文書》(A Collection of Archival,Documents From Kavalan and Siraya)。台北：國立台灣大學人類學系出版.159pp. [Besides Siraya, there are three Makatao (?) documents in roman letters.]
- 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: 1+166pp+1map. (東京大学文学部研究報告).
- . 1985. `Kulon: Yet another Austronesian language in Taiwan?' Bulletin of the Institute of Ethnology, Academia Sinica 60: 1-59.

- . 1992. 〈平埔族各語言研究瑣記(上)(下)〉 (Miscellanies of languages of sinicized ethnic groups in Taiwan) 《臺灣史田野研究通訊》(Newsletter of Taiwan History Field Research) 22:9-22 (March 1992); 23:26-42 (June 1992).
- . 2000. 'Lexical prefixes and prefix harmony in Siraya.' In *Grammatical Analysis: Morphology, Syntax, and Semantics. Studies in Honor of Stanley Starosta* (eds. V. P. De Guzman and Byron W. Bender). pp.109-128. Honolulu: University of Hawai'i Press.
- . 1996. 'Personal Pronouns in Siraya (Formosa).' In *Reconstruction, Classification, Description; Festschrift in Honor of Isidore Dyen* (ed. by B. Nothofer), pp.231-247. Hamburg: Abera Verlag Meyer & Co.
- . 2008. 「ツォウ語阿里山方言における r 音」(The r sound in Alishan dialects of Tsou). 『高一生(矢多一生)研究』[*Studies of Yata Issei (Gao I-sheng)*] 9・10:10-23. 天理：天理大学国際文化学部下村研究室。浦安：草風館。
- Tsuchida, Shigeru. (forthcoming). 'When did the sound r disappear in the Alishan dialects of Tsou?'
- Tsuchida, Shigeru, Yukihiro Yamada, and Tsunekazu Moriguchi. 1991. *Linguistic Materials of the Formosan Sinicized Populations I: Siraya and Basai*. Tokyo: Dept. of Linguistics, University of Tokyo. ix+257pp. (Reproduced with correction in Li, Paul Jen-kuei (ed.) 李壬癸編著，1997.04。《高雄縣南島語言》，高雄縣文獻叢書系列 [Austro-nesian Languages in Kao-hsiung Province] 7. 6+iii+588pp。高雄：高雄縣政府 (Kao-hsiung: Government of Kao-hsiung Province))。
- Van der Vlis, C.J. 1842. 'Formosaansche woorden-lijst, volgens een Utrechtsch handschrift.' *Verhandelingen van het Bataviaasch Genootschap* 18:437-488。
- Weng, Jia-yin 翁佳音. 1989a. 〈二十二號單語新港文書試解〉 [A trial interpretation of a piece of monolingual Sinkang manuscript No.22]。《臺灣史田野研究通訊》(Newsletter of Taiwan History Field Research 12:27-28)。
- . 1989b. 〈二十三號新港文書與西拉雅族的姓名制考〉 [Sinkang manuscript, No.23 and the naming system of the Siraya]。《臺灣史田野研究通訊》(Newsletter of Taiwan History Field Research 13:45-47)。
- . 1990a. 〈一件單語新港文書的試解〉 [A trial interpretation of a piece of monolingual Sinkang manuscript]。《民族學研究所資料彙編》(Field Materials, Institute of Ethnology, Academia Sinica) 1:143-152。
- . 1990b. 〈八十三號新港古文書譯注〉 [Translation and notes on Sinkang manuscript

土田滋

No.83]。《臺灣史田野研究通訊》(Newsletter of Taiwan History Field Research)
15:52-54。

Wu, Guo-sheng and Weng Jia-yin 吳國聖·翁佳音. 2006. 〈新港文書研究 典契的解讀
與格式〉。《建構西拉雅研討會論文集》(葉春榮主編)pp. 97-144。台南：臺南縣政府
。