

## 從文獻資料看台灣平埔族群的語言

李壬癸  
中央研究院

本文將先說明平埔族群語言的各種文獻資料，從十七世紀荷蘭傳教士所做的記錄，到十九世紀歐美人士的零星資料，直到日治時期的有計畫的調查，都為我們留下了珍貴的語言資料。可惜有清一代二百多年間，並沒留下多少語言資料：一則中國傳統文人並不重視田野調查，二則他們也沒有適當可用的記錄工具。

根據我們所能掌握的各種平埔族群語言資料，所有的平埔族群語言可以分為這三個支群：(一) 西部平埔族群：Taokas、Babuza、Papora、Hoanya、Thao，(二) 北部 Kavalan、Basay 和南部 Siraya 平埔族群，(三) 巴宰語 (Pazih)。西部平埔族群，有二種共同的音變並不見於其他任何台灣南島語言：(1) \*n 跟 \*ŋ 合併，(2) \*s 跟 \*t 合併。其中 Taokas、Babuza、Papora 跟 Hoanya 又有另外二種共同的音變：(3) \*k 消失，(4) 語詞尾的 \*-y 消失。根據各種音韻演變規則，我們繪出這五種語言的親疏遠近樹圖。北部的 Kavalan 和 Basay，南部的 Siraya 語群，事實上也跟東部的 Amis 都有一種共同的音變：(1) \*j 跟 \*n 合併，並不見於任何其他南島語言。其中 Basay 跟 Kavalan 又有二種共同的音變：(2) \*j、\*n 又進一步跟 \*N 合併，(3) \*k 分化為 k 和 q，因此這兩種語言的關係最近。對於南部平埔族群的三種語言 (Siraya, Taivuan, Makatau)，我們也比較它們之間的異同和親疏遠近關係，並且發現荷蘭時期的聖經翻譯都是以 Taivuan 語為主，而不是 Siraya 語。

關鍵詞：平埔族、語言關係、音變、合併、消失

### 1. 前言

台灣南島語言有兩大特色：(一) 語言的差異最大，(二) 保存最多的古語特徵（請詳見李壬癸 2007，Li 2008b）。因此，每一種台灣南島語言都值得作深入的研究，包括已消失的各種平埔族語言。

南島語言曾經遍布於台灣全島各地，從平原到高山都有。語言的總數大約有二十種，其中大約有一半屬於平埔族語言。台灣的平原大部分都在西部和西南部，只有一小部分

在北部（台北盆地）、東北部（蘭陽平原）和東部（花東縱谷平原）。在西部和南部平原的各種平埔族群因為最早跟外界接觸，從明末清初以來就逐步漢化，他們傳統的語言和文化在清代後期就已消失殆盡了；只有開發較晚的埔里盆地一帶和東部縱谷平原還能保留一些。有三種平埔族語言一直到二十一世紀初都還保存：在西部平原內陸地區的巴宰語和邵語，以及東部的噶瑪蘭語；其他已消失的平埔族語言，我們或多或少都有一些文獻資料（土田滋 1991, 1992，李壬癸 1992）。

平埔族的語言從北到南主要有這幾種：Kavalan、Basay、Taokas、Papora、Babuza、Hoanya、Pazih、Thao、Siraya。較少人知的有 Kulon、Qauqaut、Taivuan、Makatau 等。

## 2. 平埔族群語言的各種文獻資料

### 2.1 荷蘭傳教士（1624-1662）

台灣南島語言最早有文字記錄的是在十七世紀中葉荷據時代。荷蘭傳教士在台南地區教會了當地的土著使用羅馬拼音來書寫他們自己的語言，把整本的《馬太福音》（Gravius 1661）翻譯成「新港方言」，也就是一般人所知的「西拉雅（Siraya）」語。另有一部《基督教信仰要旨》（Gravius 1662）。除此兩部書以外，另外還有一種資料，叫作 Utrecht Manuscript，是在荷蘭 Utrecht 大學發現的西拉雅語稿件，包括詞彙資料和一篇對談。以上這三種文獻資料記錄的是嘉南平原地區的平埔族語言，是全台灣最早、最豐富、也最重要的語言文獻資料。

西部沿海地區的平埔族群語言今日都已消失了。幸而荷蘭傳教士也為其中一種留下了珍貴的資料，那就是在彰化地區的 Favorlang 方言（屬於 Babuza 語），包括一部詞典（Happart 1650）和一部基督教義傳教語料（共有 19 種文本，見 Campbell 1896）。我們只能依賴這些文獻資料來了解早期西部沿海地區的平埔族群語言的狀況了。

此外，荷蘭傳教士教會了西拉雅人羅馬拼音以後，台南地區一些部落陸續有人用來寫契約文書，稱之為「新港文書」（Murakami 1933，李壬癸 2009），從康熙年間到嘉慶年間前後一百多年（1683-1818）所寫的那些契約文書，也就成為了解那個時期各地語言異同的主要依據。

至於明鄭和清治時期前後二百多年（1662-1895），除了若干詞彙和少數歌謠以外，並沒有留下多少語言資料。一則他們沒有適當可用的記錄工具，二則傳統的中國文人並不重視田野調查工作。

### 2.2 歐美人士（1860-1900）

清朝政府採取閉關政策，八國聯軍進攻北京之後才被迫開放一些港口，1860 年以後

才有一些歐美人士陸續到台灣來，他們所寫的遊記和有關台灣南島民族的報導、零星的詞彙記錄，都是值得我們做參考的資料。

## 2.3 日治時期（1895-1945）

對台灣南島語言做有系統的調查研究始於日治時期。日本語言學者小川尚義跟淺井惠倫都曾經在各地做過平埔族群語言的調查，人類學者伊能嘉矩也留下一些相關的記錄。還有一些業餘的人士，包括警察，有時也會為我們留下珍貴的語言資料。對於多數已經消失的各種平埔族群語言，我們大都仰賴日治時期所記錄的資料（參見 Tsuchida 1982、1985，Tsuchida et al 1991，土田滋 1991，Ogawa 2006）。

淺井於 1936-37 年，為台灣北部的平埔族群 Basay 語蒐集了近一千個詞彙（Tsuchida et al 1991）以及十幾個文本。這些都是記錄大台北地區的平埔族語最豐富、最珍貴的資料。只有透過他所記錄的文本的解讀和分析，我們才可能對該語言的語法結構系統有所認識（參見 Li 1999）。

## 3. 平埔族群的類緣關係

從類緣的關係（genetic relationships）看，這些平埔族語言可以分為三群：（一）西部的 Taokas、Papora、Babuza、Hoanya 和 Thao（Li 2003b），（二）北部的 Kavalan、Basay 和南部的 Siraya、Taivuan 和 Makatau（Li 2004），（三）在中部內陸地區的 Pazih。此外，在桃園的龜崙（Kulon）語跟賽夏語關係最密切，但僅留下 45 個單語（Tsuchida 1985）。

### 3.1 西部平埔族語言

西部五種平埔族群的語言有二種共同的音變（phonological innovations），並不見於其他台灣南島語言：

1.  $*\eta > n$ ，即  $*n$  和  $*\eta$  合併，
2.  $*s > t$ （非語詞尾），即  $*s$  和  $*t$  合併。<sup>1</sup> 例如：

PAN  $*Cali\eta$  > Tao<sup>2</sup> *sarina*, Bab *harina*, Pap *sarina*, Hoa *sarina*, Tha *larina* 「耳」

PAN  $*qaNua\eta$  > Bab *loan*, Pap *loan*, Hoa *loan*, Tha *qnuan* 「牛」

PAN  $*asu$  > Bab *atu*, Pap *hatu*, Hoa *atu*, Tha *atu* 「狗」

PAN  $*esa$  > Tao *taa-nu*, Bab *na-ta*, Pap *ta-nu*, Hoa *itta*, Tha *ta-ta* 「一」

PAN  $*Ca\eta s$  > Tao *s-am-ani*, Bab *s-um-ani*, Pap *s-m-ani*, Hoa *s-am-ani*, Tha *o-m-anit* 「哭」

<sup>1</sup> 排灣語雖也有  $*s > t$ ，但  $*s$  並不跟  $*t$  合併（何大安 1998:160）。

<sup>2</sup> 語言名稱縮寫如下：Tao = Taokas，Bab = Babuza，Pap = Papora，Hoa = Hoanya，Tha = Thao，Kav = Kavalan，Ami = Amis，Sir = Siraya，Mak = Makatau。

從上面兩條音變可以看出，西部五種平埔族群語言  $*\eta > n$  的音變（即  $*n$  和  $*\eta$  合併）完全相同；而  $*s > t$  有四種語言（沿海的四種）的音變也完全相同：在語詞中的  $*s$  才變成  $t$ ，但在語詞尾的就消失了，不過邵語（Thao）則兩種語境（語詞中和尾）都變成  $t$ （即  $*s$  和  $*t$  完全合併）。因此，邵語跟西部沿海的四種平埔族群語言仍有一些差異。這五種語言的共同祖語  $*s$  和  $*t$  合併，後來沿海的四種語言  $*s$  在語詞尾丟失，而與邵語分開。

沿海四種平埔族群語言還有兩種音變：（一） $*k > \emptyset$ ，（二） $*-y > \emptyset$ （Tsuchida 1982: 6），並不見於其他台灣南島語言，包括邵語。例如：

$*kuCu > \text{Tao } usu, \text{ Bab } ocho, \text{ Pap } u\emptyset u, \text{ Hoa } usu$  「頭蝨」

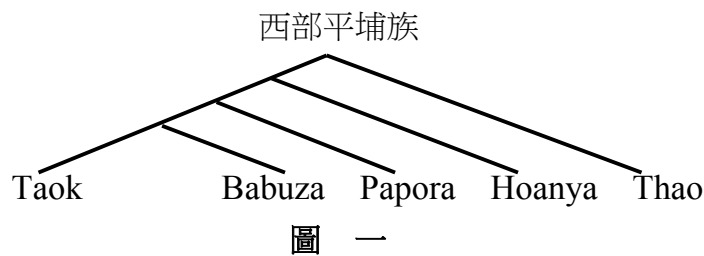
$*pirak > \text{Tao } pira, \text{ Pap } parai, \text{ Hoa } pira$  「錢」

$*babuy > \text{Bab, Pap, Hoa } babu$  「豬」

$*maCey > \text{Bab } macha, \text{ Pap } mapa, \text{ Hoa } ma\emptyset a$  「死」

可見這四種語言的關係非常近，他們分化的年代應該相當晚近，大概不到一千年。

西部五種平埔族群語言的關係如下圖所示：



### 3.2 東部語群（East Formosan）

在整個南島語系總共約有一千種語言當中，古南島語的  $*j$  和  $*n$  的合併只見於這四種台灣南島語言：巴賽（Basay）、噶瑪蘭（Kavalan）、阿美（Amis）、西拉雅（Siraya），因此這四種語言的關係最密切（Blust 1999）：

$*j$ 、 $*n > n$

$*Suaji > \text{Bas } suani, \text{Kav } suani$  「弟妹」

$*\eta ajan > \text{Bas } \eta anan, \text{Kav } na\eta an (< M),^3 \text{ Ami } \eta a\eta an (< A), \text{Sir } nanan (< M)$  「名字」

$*ajem > \text{Bas } anem, \text{Kav } anem$  「心」

$*pajey > \text{Kav } panay, \text{Ami } panay$  「稻，穀」

$*pijaH > \text{Bas } pina, \text{Ami } pina, \text{Sir } pi-pina$  「多少」

$*Sajek > \text{Kav } m-sanek, \text{Ami } ma-sanek$  「（食物將變壞）發出氣味」

值得注意的是其中三種語言屬於平埔族，而阿美卻是過去所謂的「高山族」。由此

<sup>3</sup>  $< M$  表示 metathesis「換位」， $< A$  表示 assimilation「同化」。

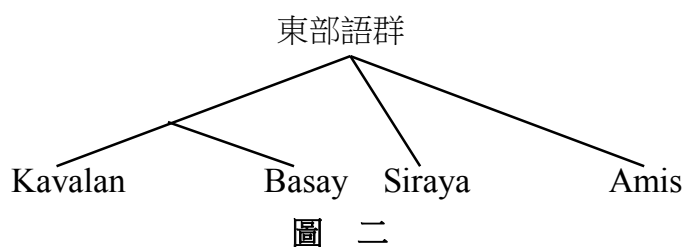
可見，平埔族和高山族之分是沒有任何學術根據的。凡是向清朝政府繳稅的就都叫作「平埔族」。其實阿美族也住在平原地區，因為開發較晚，在清代並沒有向政府繳過稅，就被歸為「高山族」了。

Blust (1999: 45) 指出，這四種語言還有兩種共同音變：(一) \*t 和 \*C 合併，(二) \*q > ʔ。可是，\*t 和 \*C 合併也見於布農語，\*q > ʔ 的語言更多，如卡語、沙語、卑南語（南王方言）、賽夏、巴宰等等。因此，這二種音變並不能證明這四種屬於同一支群。

巴賽和噶瑪蘭有兩種共同音變：(一) \*j、\*n 跟 \*N 也都合併為 n（阿美跟西拉雅並沒跟 \*N 合併），(二) \*k 分化為 k 和 q（在低元音 \*a 之前），而這種音變並不見於其他台灣南島語言。音變的證據顯示巴賽和噶瑪蘭兩種語言的關係最密切，而跟阿美、西拉雅的關係稍微疏遠一點（參見 Li 2004）。

詞形特殊偶然的變化（sporadic changes）也可提供很強有力的證據，支持巴賽跟噶瑪蘭有最密切的關係。例如，\*susu > Bas *cicu*, Kav *sisu* 「乳」（元音異化），\*piliq > Bas *p-am-ici*, Kav *p-am-il* 「挑選」（加插中綴 -am- 在第一輔音之後）或 Bas *pa-mici*, Kav *pa-mil* 「挑選」（第一輔音的鼻音取代）。

這四種台灣南島語言的關係如下圖所示：



### 3.3 巴宰語 (Pazih)

巴宰族本來在台灣中部內陸地區大甲溪和大安溪中游一帶定居，如今其傳統語言及文化全部都已消失了，只有道光年間遷至埔里盆地居住的極少數老人還保留一些母語，恐怕時日也不多了。

我 (Li 1985) 曾把巴宰跟賽夏語歸併在同一分支，而何大安 (1998: 158) 把巴宰跟泰雅併為一小分支，再跟賽夏合併為一分支，但是證據都不強。巴宰語的隸屬問題尚有待進一步研究。

## 4. 南部平埔族群語言的內部關係

根據小川 (1917) 所整理的詞彙比較表，土田 (Tsuchida *et al* 1991: 8) 指出，南部三種平埔族群語言有二條音韻演變：(一) Taivuan 的 *h* 或 *ø* 對應 Siraya 和 Makatau 的 *r*，

(二) Makatau 的 *n* 對應 Siraya 和 Taivuan 的 *l*，但有不少例外：

表一 Sirayaic Reflexes of PAn \*l and \*N 西拉雅語群對古音的反映

	PAn	Siraya	Taivuan	Makatau	
(1)	*l	r	ø ~h	r	
例	*telu	turu	too, toho	toru	「三」
	*lima	rima	hima	rima	「五」
(2)	*N	l	l	n	
例	*(qa)Nuang	luang	lowan	noang	「牛」
	*puNi	mapuli	mapuli	mapuni	「白」

根據新港文書的資料，我發現另外二條音韻演變規則：(三) Siraya 的 *s* 對應 Taivuan 和 Makatau 的 *r* 或 *d*，(四) Taivuan 元音之間的舌根音丟失：

表二 Sirayaic Reflexes of PAn \*D, \*-k- and \*-S-, \*-R-

	PAn	Siraya		Taivuan		Makatau	
		Sinkang 新港	Tohkau 卓猴	Wanli 灣里	Matau 麻豆	Lower Tamtsui 下淡水	
(3)	*D	s	s	r~d	r~d	r~d	
(4)	*-k-	-k-	-k-	ø	ø	-k~ø	
	*-S-, *-R-	-g~-h-	-g-	ø	ø	--	
(3b)	*Daya	saija	saija	raija	--	--	「東」
	*lahuD	raos	raos	raur	--	--	「西」
	*DapaN	sapal	--	rapan	--	--	「腳」
		sa	sa	ra, da	ra, da	ra, da	「和」
		hisō	hisō	hairo, ro	haijro, do	--	「若」
		posoh	--	poroh	--	--	「地」
		maisisang	--	--	--	maeraerang	「大官」
(4b)		akosaij	akusiwo	ausaij	ausaij	akusai	「無」
		tarokaij	--	taroej	--	tarauwei	「人名」
	*DuSa	soso(h)a	--	--	--	--	「二」
	*baqeRu	vaho	--	--	--	--	「新」
		dagogh	dagogh	daoh	daoh	--	「價」
		ligig	--	liih	--	--	「砂」
		matagi-vohak	--	mataij-vohak	--	--	「悔」

綜合各種文獻資料，南部三種平埔族語言共有四條音韻演變規則：

表三 Summary of Sirayaic Reflexes of PAn \*l, \*N, \*D, \*-k- and \*-S-, \*-R-

	PAn	Siraya	Taivuan	Makatau
(1)	*l	r	ø~h	r
(2)	*N	l	l	n
(3)	*D, *d	s	r~d	r~d
(4)	*-k-	-k-	ø	-k~ø
	*-S-, *-R-	-g-	ø	--

除音韻演變不同之外，表示未來的後綴 Taivuan 的 *-ah* 也跟 Siraya 的 *-ali* 和 Makatau 的 *-ani* 有明顯的不同。從音變和構詞兩方面來觀察，Taivuan 跟其他兩種語言似乎有較多的差異。這三種語言之間的關係，暫如下圖所示：

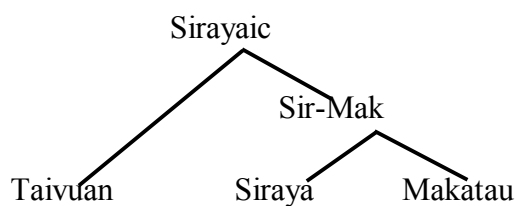


圖 三

十七世紀荷蘭傳教士所記錄的三種文獻資料各為什麼語言？試比較它們所顯示的語音對應關係，請見下頁表四。

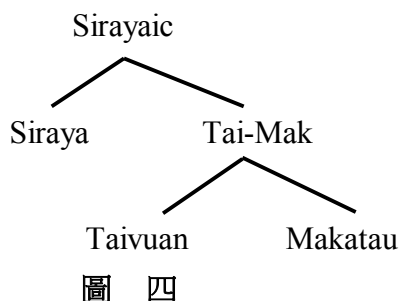
表四 Utrecht, Matthew and Formulary 三種文獻資料的比較

	Utrecht Ms 烏忒稿本	St. Matthew 馬太福音	Formulary 基督教義	
(5)	*D	s	r, d	r, d
	*Daya	taga-seia	reia	-- 「東」
	*lahuD	taga-raos	raour	-- 「西」
	*DapaN	sapal	rahpal	rahpal 「腳」
	*DuSa	so-soa	dou-rouha	rou-rouha 「二」
	*likuD	ricos	rikour	rikour 「背」
		sama	rama	dama 「父」
		sa	ra	ra 「但是」
		soo, sou	rou	dou, rou 「若, 當」
		isang	irang	irang 「大」
		sasim	rarim	mou-rarim 「下」
		pesanach	paeraenaeh	paeraenaeh 「樹」
		massou	--	marou 「玉米」
		ka-pousoch-ang	pourough	pourogh 「地」
(6)	*-k-	-k-, ø	-k-, ø	ø
		acoussey	akousi, aousi	aoussi 「無」
	*(i)aku	iau, -au	jau, -au	jau 「我」
(7)	*-S-/*-R-	-h-, ø/-g-, -h-	-h-, ø/-h-, ', ø	-h-, ø/h, ', ø
	*kaSu	cau [kaw]	kow	kow 「你」
	*DuSa	so-soa	dou-rouha	rou-rouha 「二」
	*CaSiq	t-m-ahy	--	-- 「縫」
	*waRi	wagi	wæ'i	wæ'i 「日」
	*wiRi	ougi	u-i	ou-i 「左」
	*baqəRu	vaho	vahæu, va'æu	vahæu, va'æu 「新」
	*kaRaC	k-m-agat	--	-- 「咬」
		ligig	li'igh	-- 「砂」
		ma-dagoa	--	dæeua 「總」
(8)		-a, -al, -ale	-ah, -auh	-a, -ah, -al 「未來」



表四資料顯示：Utrecht 稿本是根據 Siraya 語，而《馬太福音》跟《基督教義》那二部書卻都是根據 Taivuan 語。一般人都誤以為它們都是西拉雅語。

前面根據音韻演變規律的多寡，我曾暫時認為 Taivuan 語大概先分化出來。可是若考慮音韻演變時代的先後，就會有不同的結論。上面表三第 3 條的音變 PAN \*D > Siraya s，Taivuan 和 Makatau *r* ~ *d* 是在荷蘭人來台之前就已完成，是最早的音變。第 4 條，元音之間舌根音的丟失是在十七世紀上半葉才開始，因為只有部分的詞彙消失，而有些詞彙仍然保存舌根音，音變的時代較晚。第 1 條，\*l 變成 *h* 或丟失的時代更晚，因為荷蘭時代聖經的翻譯資料都保存，在十八世紀灣里社的文書（1770 年）和麻豆的文書（1781 年）也都還保存，直到十九世紀末所記錄的資料才丟失。因此，從音變時代的先後來看，似以 Siraya 最先分化出來為宜：



## 5. 討論和結語

現存的三種平埔族語言：噶瑪蘭、巴宰、邵，本人都親自調查研究了三十多年了，對這些語言現象的掌握自然遠勝於其他已消失的各種平埔族語言。對於已消失的各種平埔族語言，我們必須透過有限的文獻資料進行解讀、分析和研究，對台灣南島語言的歷史才可能有較全面的掌握。

儘管絕大多數平埔族語言的資料並不多，只要我們用心去搜尋、整理、分析，仍然可以得到不少有價值的資訊。已消失的各種平埔族語言只有三種有文本資料：南部的西拉雅語群、西部的法佛朗語（Favorlang）、北部的巴賽語，因此，對於這些語言的語法結構系統我們才有可能進行分析研究。這三種語言的地理分布還相當不錯，並非集中在某一個小區域；而且它們分別隸屬於兩種不同的支群：法佛朗屬於西北支群，而西拉雅跟巴賽屬於東部支群（參見 Blust 1999）。其他已消失的幾種平埔族語言，我們都只有若干詞彙資料，最少的只有十個詞，最多的也只不過三百多個詞（參見李壬癸 1992）。

中部內陸地區的邵族，在地理上也跟它們（貓霧揀(Babuzá)、洪雅(Hoanya)）相鄰接，語言的關係也還相當接近。這五種平埔族大概在最近二千年內才逐步分化，依序為邵、洪雅、巴布拉（Papora）、貓霧揀與道卡斯（Taokas），由南向北擴散。可是，同樣在

中部內陸地區的巴宰語，卻跟這些語言的關係相當疏遠；它的隸屬關係仍然有待進一步研究才能確定。巴宰族分化出來的時代應當更早，大概超過三千年。

由最靠北部的巴賽和噶瑪蘭到南部的西拉雅語群，以至到東部的阿美，這四種語言的關係最接近，可是地理上卻離得最遠。可見語言的隸屬關係跟地理位置並沒有一定的關聯性。即使如此，同樣在北部的巴賽語和噶瑪蘭語彼此之間的關係確實也最接近（Li 2004）。這個支群的擴散大概是經由海路，而不是經由陸路，並且是數千年以前的事了。有關各族群的擴散史，請參見 Li 2008a。

本文把討論的重點放在西部平埔族支群和「東台」支群，而略去較少人知道的猴猴（Qauqaut）族和龜崙（Kulon）族。其主要原因，一則語言資料實在太少（猴猴只有 10 個單字，龜崙只有 45 個單字），二則關於這兩個族群過去已有專文討論過（關於龜崙請參見 Tsuchida 1985，關於猴猴請參見馬淵 1931 等）。

## 引用文獻

- Blust, Robert. 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun & Paul Jen-kuei Li, 31-94. Symposium Series of the Institute of Linguistics (Preparatory Office), No.1. Taipei: Academia Sinica.
- Campbell, Rev. William. 1888. *The Gospel of Formosan (Sinkang Dialect), with Corresponding Versions in Dutch and English Edited from Gravius's Edition of 1661*. London: Trübner and Co.
- \_\_\_\_\_. 1896. *The Articles of Christian Instruction in Favorlang Formosan, Dutch and English from Vertrecht's Manuscript of 1650*. London: Kegan Paul, Trench, Trübner and Co.
- 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 Hartoch Boekverkoper, inde Oude Hoogstraat, inde Boeck en Papier-winckel.
- \_\_\_\_\_. 1662. *Patar Ki Tna-'msing-an Ki Christang, ka Taukipapatar-en-ato tmaeu'ug tou Sou KA MAKKA-SIDEIA, 't Formulier des Christendoms Met de Verklaringen van dien, Inde Sideis-Formosaansche Tale*. Amsterdam: Michiel Hartogh.
- Happart, Gilbertus. 1650. *Woord-boek der Favorlangsche taal, waarin het Favorlangs voor, het*

- Duits achter gestelt is. Reproduced in W. R. van Hoëvell (1842), *Verhandelingen van het Bataviaasch Genootschap* 18: 382-430.
- Li, Paul Jen-kuei (李壬癸). 1985. The position of Atayal in the Austronesian family. *Austronesian Linguistics at the 15th Pacific Science Congress*, ed. by Andrew Pawley & Lois Carrington, 257-280. Pacific Linguistics C-88.
- \_\_\_\_\_. 1999. Some problems in the Basay language. *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul Li, 635-664. Symposium Series of the Institute of Linguistics (Preparatory Office), No.1. Taipei: Academia Sinica.
- \_\_\_\_\_. 2003a. Notes on Favorlang, an extinct Formosan language. *English- Favorlang Vocabulary by Naoyoshi Ogawa*, ed. by Paul Li, 1-13. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Asia- African Lexicon Series No. 43. A revised version appeared in *POLA Forever: Festschrift in Honor of Professor William S.-Y. Wang on His 70th Birthday*, ed. by Dah-an Ho and Ovid J. L. Tzeng, 175-194. Taipei: Institute of Linguistics, Academia Sinica.
- \_\_\_\_\_. 2003b. The internal relationships of six western plains languages. *Bulletin of the Department of Anthropology* 61: 39-51.
- \_\_\_\_\_. 2004. Origins of the East Formosan peoples: Basay, Kavalan, Amis and Siraya. *Language and Linguistics* 5. 2: 363-376.
- \_\_\_\_\_. 2008a. Time perspective of Formosan aborigines. *Past Human Migrations in East Asia and Taiwan: Matching Archaeology Linguistics and Genetics*, ed. by Alicia Sanchez-Mazas, Roger Blench, Malcolm Ross, Ilia Peiros and Marie Lin, 211-218. London and New York: Routledge Curzon.
- \_\_\_\_\_. 2008b. The great diversity of Formosan languages. *Language and Linguistics* 9. 3: 523-546.
- \_\_\_\_\_. 2009. Linguistic differences among Siraya, Taivuan and Makatau. *Austronesian Historical Linguistics and Culture History, a Festschrift for Robert A. Blust*, ed. by Alexander Adelaar and Andrew Pawley, 387-396. Pacific Linguistics 601.
- Li, Paul J. K. and Shigeru Tsuchida (李壬癸, 土田滋). 2001. *Pazih Dictionary* (《巴宰語詞典》). Institute of Linguistics (Preparatory Office), Academia Sinica Monograph Series No. A2. Taipei: Academia Sinica.
- \_\_\_\_\_. 2006. *Kavalan Dictionary*(《噶瑪蘭語詞典》). Institute of Linguistics, Academia

- Sinica Monograph Series No. A19. Taipei: Academia Sinica.
- Murakami, Naojirô (村上直次郎). 1933. 《新港文書》. 台北帝國大學文政學部紀要第二卷第一號. Formosa: Taihoku Imperial University.
- Ogawa, Naoyoshi (小川尚義). 1917. Siraia, Makatao, Taivoan. Unpublished manuscripts. (kept at Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies).
- \_\_\_\_\_. 2006. 《台灣蕃語蒐錄》 *A Comparative Vocabulary of Formosan Languages and Dialects*, ed. by Paul Jen-kuei Li and Masayuki Toyoshima. Asian and African Lexicon Series No.49. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Tsuchida, Shigeru (土田滋). 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.
- \_\_\_\_\_. 1985. Kulon: Yet another Austronesian language in Taiwan? *Bulletin of the Institute of Ethnology* 60: 1-59.
- Tsuchida, Shigeru, Yukihiro Yamada and Tsunekazu Moriguchi (土田滋, 山田幸宏, 森口恆一). 1991. 《台灣・平埔族の言語資料の整理と分析》 (Linguistic Materials of the Formosan Sinicized Populations I: Siraya and Basai). 東京：東京大學(University of Tokyo).
- Utrecht Manuscripts. 1650?. (烏忒稿本) Utrecht University Library.
- 土田滋. 1991. 〈平埔族諸語研究雜記〉，《東京大學言語學論集》，12: 146-179。
- \_\_\_\_\_. 1992. 〈平埔族諸語研究雜記〉，《中央研究院臺灣史田野研究通訊》，22: 9-22，23: 26-42。
- 李壬癸. 1992. 〈台灣平埔族的種類及其相互關係〉，《臺灣風物》，42. 1: 211-238。
- \_\_\_\_\_. 2001. 〈邵族的地位—兼評白樂思 (Blust 1996) 的邵族地位說〉，詹素娟、潘英海編《平埔族群與臺灣歷史文化論文集》，165-184。台北：中央研究院臺灣史研究所籌備處。
- \_\_\_\_\_. 2007. 〈台灣南島語言的回顧和展望〉，《原住民族語言發展論叢：理論與實務》，1-11。台北：行政院原住民族委員會。
- \_\_\_\_\_. 2009. 《新港文書研究》，《語言暨語言學》專刊甲種之 34。台北：中央研究院語言學研究所。
- 何大安. 1998. 〈台灣南島語的語言關係〉，《漢學研究》，16. 2: 141-171。

馬淵東一. 1931. 〈研海地方に於ける先住民の話〉，《南方土俗》，1. 3。

---

李壬癸

中央研究院語言學研究所

paulli@gate.sinica.edu.tw

## Plain Languages as Seen from Written Documents

Paul Jen-Kuei LI

*Academia Sinica*

Many Formosan languages were spoken in the Taiwan plains; most are now extinct. For data on them, we must rely on the writings of 17<sup>th</sup> century Dutch missionaries, of early 20<sup>th</sup> century Japanese field workers, and of a few Chinese or other Western observers from the 17<sup>th</sup> through 19<sup>th</sup> centuries. For Siraya in the southwestern plains and Favorlang on the central west coast, there is a fair amount of documentation. For all the other western plain languages, however, we have less than 400 lexical items. We still have informants for only three plain languages: Pazih, Thao, and Kavalan. For these, language data and linguistic information are much more abundant and reliable.

All the plain languages can be classified into three main subgroups based on phonological innovations: (1) Western Plains: Taokas, Papora, Babuza, Hoanya, and Thao; (2) the so-called “East Formosan”: Basay, Kavalan, Siraya, and Amis; and (3) Pazih. The five western plain languages share exclusively two phonological innovations: (1) merger of Proto-Austronesian \*n and \*ng, and (2) merger of \*s and \*t. The western plain languages, except for Thao, lost word-final \*-s. Thao is, therefore, somewhat less closely related to the other four languages. Proto-Austronesian \*j and \*n have merged in the four East Formosan languages, a merger not found anywhere else. Basay and Kavalan share two innovations, indicating a closer relationship: (1) merger of \*j, \*n and \*N, and (2) split of \*k into *k* and *q* (\*\_a), so they are more closely related to each other.

It is interesting to note not only that geographically adjacent languages may be closely related, as is the case with the five western plain languages, but also that geographically distant languages can likewise be closely related, as in the case of the four so-called “East Formosan” languages.

In Li 2009, I examine the language data of the southwestern plains in greater detail and sort out the linguistic differences among Siraya, Taivuan, and Makatau. As it turns out, the Dutch biblical translations are based on Taivuan rather than Siraya.

Key words: plain tribes, language relationships, phonological innovations, merger, loss