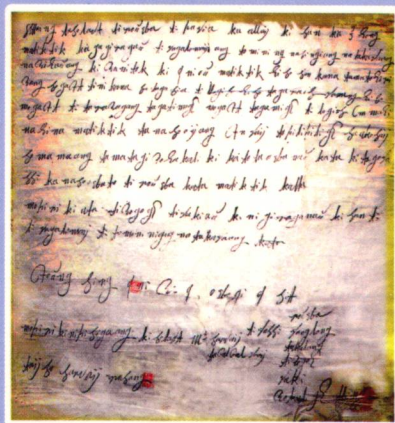


《語言暨語言學》專刊系列之三十九

新港文書研究

李壬癸／編著

黃秀敏 簡靜雯 陳鳴鸞 許淑鈴／助理



中央研究院 語言學研究所

2010

H284
20121

2010-1-039-001-000010-1

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES 39

《語言暨語言學》專刊系列之三十九

Studies of Sinkang Manuscripts

新港文書研究

by

Paul Jen-kuei Li

李壬癸／編著

with the assistance of Hsiu-min Huang, Chin-wen Chien,
Amy Min-luan Chen & Shu-ling Hsu

助理

黃秀敏 簡靜雯 陳鳴鸞 許淑鈴



Institute of Linguistics, Academia Sinica, Taipei, Taiwan

中央研究院 語言學研究所

2010

LANGUAGE AND LINGUISTICS

LANGUAGE AND LINGUISTICS is a department of the INSTITUTE OF LINGUISTICS, ACADEMIA SINICA. It furthers the Institute's objective of excellence in research, scholarship, and education by publishing pioneering articles and monographs in Linguistics and other related fields. All publications have been reviewed rigorously according to academic standards.

Copyright © 2010 by INSTITUTE OF LINGUISTICS, ACADEMIA SINICA
128, Sec. 2, Academia Road, Nankang, Taipei 11529, Taiwan, R.O.C.

<http://www.ling.sinica.edu.tw>

All rights reserved.

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES 39

By Paul Jen-kuei Li

With the assistance of Hsiu-min Huang, Chin-wen Chien, Amy Min-luan Chen & Shu-ling Hsu

Studies of Sinkang Manuscripts

《語言暨語言學》

《語言暨語言學》隸屬於中央研究院語言學研究所，以出版語言學及相關領域之最新研究成果為宗旨。所有出版品均經過嚴格學術審查。

《語言暨語言學》專刊系列之三十九

新港文書研究

李壬癸 編著

黃秀敏 簡靜雯 陳鳴鸞 許淑鈴／助理

出版：中央研究院語言學研究所

發行：中央研究院語言學研究所

115 台北市 南港區 研究院路 2 段 128 號

<http://www.ling.sinica.edu.tw>

版權所有 翻印必究

印刷：文匯印刷資訊處理有限公司

初版：中華民國九十九年六月 [2010]

定價：新台幣 1,200 元／US\$80

ISBN-13: 978-986-02-3342-1

GPN: 1009901571

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES

EDITORIAL BOARD

<i>Editor in Chief</i>	Chih-Chen Jane Tang (Academia Sinica)
<i>Honorary Editor</i>	Jiun-Shiung Wu (National Chiayi University)
	President, Linguistic Society of Taiwan
<i>Executive Editor</i>	Elizabeth Zeitoun (Academia Sinica)

Editorial Board Members

- Mark C. Baker (Rutgers University)
- Robert A. Blust (University of Hawai'i, Manoa)
- Gennaro Chierchia (Harvard University)
- Sandra Chung (University of California, Santa Cruz)
- Bernard Comrie (Max Planck Institute for Evolutionary Anthropology)
- William Croft (University of New Mexico)
- Hiroya Fujisaki (University of Tokyo)
- C.-T. James Huang (Harvard University)
- Larry M. Hyman (University of California, Berkeley)
- Christine Lamarre (Institut National des Langues et Civilisations Orientales)
- Richard K. Larson (Stony Brook University)
- Paul Jen-kuei Li (Academia Sinica)
- Ian Maddieson (University of California, Berkeley)
- Alec Marantz (New York University)
- James A. Matisoff (University of California, Berkeley)
- Alain Peyraube (Ecole des Hautes Etudes en Sciences Sociales)
- Elisabeth Selkirk (University of Massachusetts, Amherst)
- Hsiu-Fang Yang (National Taiwan University)

Editorial Assistants

Chun-yu Kuo (Academia Sinica)
Vicky Y-K. Chen (Academia Sinica)
Hsin-Hui Hsu (Academia Sinica)

INTRODUCTION

Arrangement of this Monograph

This volume is divided into three main parts.¹ Part 1 contains all 101 of the Sinkang manuscripts that appeared in Murakami (1933). Part 2 contains a further 69 Sinkang manuscripts collected since then. And Part 3 consists of photocopies of all available Sinkang manuscripts, 150 in total. We follow the same order as in Murakami, including the manuscripts collected from the villages of Sinkang, Tohkau, Matau, Taibulang, Lower Tamsui, and Katin, and in that order. The recently uncovered Sinkang manuscripts mostly come from the village of Sinkang, four from Tohkau (Siraya), four from Matau (Taivuan), and twelve from Wanli (Taivuan); these latter all being bilingual are consequently of great value.² Unfortunately, we have uncovered nothing new from Makatau villages; Murakami himself had collected only four manuscripts there. This fact seems to indicate that the practice of writing in the native language was never well developed in Kaohsiung or Pintung in the deep south. We use the term "Sinkang manuscripts" in a broad sense; i.e. to encompass all manuscripts written in native languages formerly spoken on the southwestern plains of Taiwan.

The map shows the geographical distribution of Siraya, Taivuan, and Makatau villages on the southwestern plains of Taiwan by the end of the nineteenth century (see p.13).

A list of the Sinkang documents includes for each a manuscript number assigned by Ogawa and Murakami, village where it was written, date, the original document holder, and a brief description of the manuscript (whether original or reproduced, mono- or bilingual, contract or account, with or without an official chop). Dating of most manuscripts can be determined from references to Ching Dynasty emperors, which can then be related to Gregorian years. The 170 manuscripts are arranged in chronological order in a table on

¹ I am indebted to Ogawa's pioneer studies of Sinkang manuscripts and to my assistants, Hsiu-min Huang, Chin-wen Chien, Amy Min-luan Chen, and Shu-ling Hsu. The work of this volume was supported in part with grants from the Academia Sinica, National Science Council (NSC 95-2411-H-001-010-H, NSC 97-2410-H-001-065-MY3) and the Chiang Ching-kuo Foundation for International Scholarly Exchange (RG001-D-07).

² In a bilingual manuscript, the content of the text in the native language may not be identical to the Chinese.

pages 16-22, except for nine at the end which cannot be dated.

This is a volume of texts with translation. It is a challenging job indeed to try deciphering these Sinkang manuscripts without benefit of a living language consultant. Ogawa's unpublished manuscripts are very helpful, and we have utilized them all. Word-for-word interlinear English glosses³ and free Chinese translations are given for each text when possible. Only 34 texts have free translations in both Chinese and English. Despite our best efforts, we are still unable to identify some lexical forms or translate some sentences, thus leaving gaps in our work.

Appendix 1 is a list of Siraya names occurring in the manuscripts. A name in Siraya can be a given name or a family name, and in that order when both appear. A given name can be male or female, and they are kept distinct.

Linguistic Differences among Siraya, Taivuan, and Makatau⁴

1. Introduction

Siraya, Taivuan, and Makatau were Formosan languages or dialects formerly spoken in the southwestern plains of Taiwan. Roughly speaking, Siraya was in the coastal area of the Tainan Plain and Taivuan mostly inland on the Tainan Plain to the north, while Makatau was in Kaohsiung and Pingtung prefectures to the south (see map on p.13). The languages or dialects probably became extinct in the first half of the nineteenth century (Li 2002). Dutch missionaries left behind three main written documents, namely *The Gospel of St. Matthew in Formosan Sinkang Dialect* (Gravius 1661, henceforth *St. Matthew*), *'t Formulier des Christendoms* (Gravius 1662, henceforth *Formulary*), and the *Utrecht Manuscript* (unknown author, published in Van der Vlis 1842).⁵ Ever since then only short wordlists were recorded in various villages of the southwest plains at different times between 1717 and 1917 by the Chinese, Europeans, and Japanese. Ogawa (1917) assembled these wordlists and classified them into

³ If they are loan words from Chinese, then we will simply give the original Chinese characters instead of English glosses, e.g., *ni* 年 'year', *ka* 甲 'unit of land', *haptang* 合同 'contract', *chianliong* 乾隆 'title of an emperor'.

⁴ This part appeared as a paper with the same title in Alexander Adelaar and Andrew Pawley (2009), eds., *Austronesian Historical Linguistics and Culture History: A Festschrift for Robert Blust*. In preparing this paper, I have benefited from Ogawa's pioneer work on Siraya, and from Tsuchida and Adelaar's valuable suggestions.

⁵ Adelaar (1997:364ff) also discusses dialect variations between the *Utrecht Manuscript* and the *St. Matthew/Formulier*.

three main groups: Siraya, Makatau, and Taivuan, and in that order. There are altogether 75 villages or sources of language data and 163 lexical entries in his comparative wordlist (see Tsuchida et al. 1991). Due to the paucity of language data in that area, his comparative wordlist is extremely valuable, especially for Makatau. A map in Tsuchida et al. (1991:ix) shows the location of 39 villages, using three different symbols to mark the three different groups of languages or dialects, thus giving an idea as to the geographical distribution of these vanished linguistic communities on the southwestern plains.

Based on Ogawa's comparative wordlist, Tsuchida pointed out that the three groups have different reflexes of PAn *l and *N, as shown below:

Table 1: Sirayaic Reflexes of PAn *l and *N

	PAn	Siraya	Taivuan	Makatau	
(1)	*l	r	ø~h	r	
e.g.	*telu	turu	too, toho	toru	'three'
	*lima	rima	hima	rima	'five'
(2)	*N	l	l	n	
e.g.	*(qa)Nuang	luang	lowan	noang	'cow'
	*puNi	mapuli	mapuli	mapuni	'white'

But there are many exceptions to the rules, as Tsuchida was aware of, perhaps due to the poor or inaccurate transcriptions of the language data from various sources. If, however, these correspondences were unproblematical, then each group would differ from the other two by only one phonological innovation. These might be regarded as dialectal differences as commonly found in other Formosan languages such as Rukai (Li 1977) or Atayal (Li 1981).

2. The Evidence from the Sinkang Manuscripts

Aside from Dutch missionary documents and short wordlists from the southwestern plains, we have at our disposal a third type of language data, the so-called "Sinkang manuscripts", which were contracts written in a roman alphabet. These texts were found in various villages in the southwestern plains that belonged to the three different groups. For example, Sinkang, Tohkau, and Kongana belong to Siraya; Wanli, Matau (further discussion later), and Taibulang belong to Taivuan; Lower Tamtsui and Katin belong to Makatau. The earliest text is dated 1663, or more reliably 1683, and the latest 1818.

Murakami (1933) collected 101 manuscripts.⁶ My colleagues and I have collected 170 manuscripts.⁷ The great majority came from Siraya villages, only 23 came from Taivuan villages, and 4 from Makatau villages. Among these, 69 were written in both Chinese and a native language, while the remaining are monolingual. A careful study of these texts may reveal significant linguistic differences, not found in simple wordlists.

My assistants, Hsiu-min Huang and Chin-wen Chien, and I have worked on Sinkang manuscripts for eight years (2001-2009). Although they are extremely hard to interpret, we have tried to decipher and transcribe them all, determine word and sentence boundaries, identify each lexical item, and give interlinear glosses and free translations whenever possible. All 170 texts have now been digitized in computer files. These texts have revealed some interesting facts about the languages or dialects of the southwestern plains.

2.1 Phonological Evidence

In addition to the two phonological innovations observed by Tsuchida et al. (1991), we can add two more based on the Sinkang data, as illustrated below:

⁶ Of the 101 manuscripts collected by Murakami, 87 are from the village of Sinkang, 6 from Matau, 3 from Tohkau, 1 from Taibulang, 1 from Lower Tamtsui, and 3 from Katin. The 87 Sinkang manuscripts are treated as the main body of his monograph, while all the other 14 are given in the appendices.

⁷ Of the 170 manuscripts, 1 came from Backloan (Siraya, not in Murakami), 7 from Tohkau, 10 from Matau, 1 from Taibulang (Taivuan), 12 from Wanli (Taivuan, not in Murakami), 1 from Lower Tamtsui, 3 from Katin, and 2 from Gutiaupo 牛稠埔 (Siraya, not in Murakami).

Table 2: Sirayaic Reflexes of PAN *D, *-k- and *-S-/*-R-

	PAn	Siraya		Taivuan		Makatau	
		Sinkang	Tohkau	Wanli	Matau	Lower Tamtsui	
(3)	*D	s	s	r~d	r~d	r~d	
(4)	*-k-	-k-	-k-	ø	ø	-k-~ø	
	*-S-, *-R- ⁸	-g~-h-	-g-	ø	ø	—	
(3b)	*Daya	saija	saija	raiya	—	—	‘east’
	*DapaN	sapal	—	rapan	—	—	‘foot’
	*lahuD	raos	raos	raur	—	—	‘west’
		sa	sa	ra, da	ra, da	ra, da	‘and’
		hiso	hiso	hairo, ro	haijro, do	—	‘if, as’
		posoh	—	poroh	—	—	‘land’
		maisasang	—	—	—	maeraerang	‘magistrate’
(4b)		akosaij	akusiuo	ausaij	ausaij	akusai ⁹	‘have not’
		tarokaij	—	taroej	—	tarauwei ¹⁰	‘name’
	*DuSa	soso(h)a	—	—	—	—	‘two’
	*baqeRu	vaho	—	—	—	—	‘new’
		dagogh	dagogh	daoh	daoh	—	‘price’
		ligig	—	liih	—	—	‘sand’
		matagi-	—	mataij-	—	—	‘regret’
		vohak		vohak			

As shown in the examples above, *s* in Siraya corresponds to *r* or *d* in Taivuan and Makatau in word-initial or final position, derived from PAN *D or *d, as illustrated in (3). As shown in (4), *k* or *g~h* in Siraya is lost in Taivuan in word-medial position. The *k* in Siraya is derived from PAN *k, and the *g~h*, which is interpreted as velar fricative *x* by Adelaar (1999), is derived from PAN *S or *R (Adelaar 1999:334).

⁸ [x] (graphemes *g*, *gh*, or *h*) is historically derived from PAN *S or *R; see Table 6 below for more examples. There is extremely limited vocabulary in all Sinkang manuscripts, and I can identify few certain PAN cognates for such a derivation in these Sinkang manuscripts, such as *sosoha* or *sosoa* 'two' and *v(ah)ho* 'new' found only in Sinkang.

⁹ Although the term 'not have' is unavailable for Lower Tamtsui, the form *akusai* is cited for Taikiattian, another village of Makatau in Ogawa's list, and the form *akosaij* appears in a manuscript from Gutiaupo, which might be another village of Makatau or Siraya (see Tsai 2002, Appendix 1, p.3).

¹⁰ The personal name *tarauwei* appears in a Katin text (Murakami 1933:144). This shows that *-k-* is occasionally lost in certain Makatau dialects.

Rule (3) shows Siraya in contrast with Taivuan and Makatau, while Rule (1) shows Taivuan in contrast with Siraya and Makatau. Rule (4) shows that the medial velar obstruents *k* and *g* [x] are lost in Taivuan, but retained in both Siraya and Makatau.¹¹ It is an innovation in Taivuan.

Matau was thought to belong to the Siraya group by Chinese and Japanese scholars, as indicated in Ogawa's (1917) grouping and a map in Tsuchida et al. (1991:ix). However, the phonological innovations in Matau generally indicate that it should belong to the Taivuan group rather than to Siraya:

Matau *l > zero or *h*, e.g. *telu > *tao* 'three', *lima > *hima* 'five'; and *D > *r, d* (see Table 2 above).

Both Matau and Wanli villages of Taivuan are in the coastal and transitional area, geographically close to Sinkang and Tohkau villages of Siraya. The phonological differences, especially (3) and (4), between Siraya, Taivuan, and Makatau, are quite regular.

2.2 Morphological Evidence

In addition to the phonological differences, a type of morphological difference can be observed, as shown below:

Table 3: Sirayaic future markers

Siraya		Taivuan		Makatau	
Sinkang	Tohkau	Wanli	Matau	Lower Tamtsui	
-ali, -ili	-ati, -ili	-ah	-ah	-ani	'future'

The verbal suffix indicating future also shows that Taivuan differs from both Siraya and Makatau. Let it be noted that *-ani* in Makatau regularly corresponds with *-ali* in Siraya. Again, this bit of morphological evidence also indicates that Matau is Taivuan instead of Siraya.

¹¹ Ogawa investigated the Piathau dialect of Siraya in 1921. But he cited the form *li'igh* 'sand' for Siraya dialects (Ogawa 2006:354), probably taken from *The Gospel of St. Matthew*; see Table 6 in §3 below.

2.3 Lexical Evidence.

Tsuchida et al. (1991:7-8) pointed out the following lexical differences in the three groups, as based on Ogawa's comparative wordlist:

Table 4: Lexical differences in Sirayaic

(5)	Siraya	Taivuan	Makatau	
	siraya	taivoan	makatao	'aborigines (self-appellation)'
	it	tau	lihu	'liquor, wine'
	pak	harau	buka	'rice'

Nevertheless, lexical borrowing between the groups is often attested. The term for 'aborigines' in Matau is *siraiya* instead of the anticipated form *taivoan*. The term for 'wine' in Kongana (Siraya) is *dihō* instead of the anticipated form *it*; see Texts 5 and 11 published in Li (2002). Similarly the term for 'wine' in *St. Matthew* is also *dihou*. The term for 'rice' in Laupi (Makatau) is *pak* instead of the anticipated form *buka*. In short, unlike phonological innovations, lexical evidence is not very reliable for language subgrouping.

2.4 Summary

For the sake of convenience, let's summarize what we have discussed so far, in Table 5:

Table 5: 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	—

In addition, Taivuan has a suffix *-ah* indicating 'future' different from both Siraya *-ali* and Makatau *-ani*.

Two phonological innovations, (*l > zero or h and loss of medial consonants *-k-, *-S- and -R-), plus one piece of morphological evidence, show that Taivuan is in maximal contrast with Siraya and Makatau. Makatau differs from both Siraya and Taivuan only in Rule (2), which is not significant. It shares Rules (1), (4), and the correspondence seen in the Sirayaic future markers with Siraya, but it shares only Rule (3) with Taivuan. As a preliminary conclusion, it seems that Taivuan constitutes a first split from the Sirayaic group, while Siraya and Makatau are more closely related, as shown below:¹²

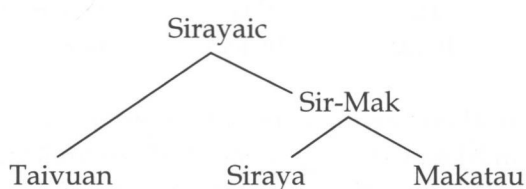


Figure 1: Sirayaic stemma based on numbers of shared phonological innovations

It seems reasonable to say that Taivuan is a separate language, but it is not clear whether Siraya and Makatau can be treated as separate languages until we find greater linguistic difference.

3. The Dutch Missionary Documents

Is *St. Matthew* based on Siraya or Taivuan? The following vocabulary found in *St. Matthew* seems to indicate that it is Taivuan:

- (6) *D, *Z > r, d, cf. (3) above

*Daya > *reia* 'east', *lahuD > *raour* 'west', *DapaN > *rahpal* 'foot', *likuD > *rikour* 'back', *ZaNum > *raloum* 'water', *da* 'and' (*sa* in Sinkang), *pourough* 'land' (*posoh* in Sinkang), *haouroung* 'to steal' (*haosong* in Sinkang), *laumari* 'coins' (*lomasi* in Sinkang), *ka-harim-auh* 'will be forgiven' (*ka-hasim-ing* in Sinkang)

¹² Adelaar's interpretation (p.c.) of Table 5 is that there is no clear subgrouping pattern emerging. He also notes that in Rule (2) in Table 5, Makatau *n* would be an innovation, rather than a retention, if PAn is reconstructed as *L.

- (7) Loss of *-k-* and *-g-* [x] < **-S-*, cf. (4) above

Loss of *-k-*: *(i)aku > jau, *-au* '1 sg'; *aousi*, *akousi* 'not have'

The loss of intervocalic *-k-* seems to be restricted to a few special grammatical categories, namely personal pronoun and negative, and note the free variation in *aousi* ~ *akousi*. It does not apply to ordinary vocabulary, e.g. *bukeS > voukugh 'hair', *likuD > rikour 'back' and *takut > takout 'to fear'.

Loss of *-g-*: *daæugh*, *daæuh* 'price' (*dagogh* in Sinkang); *li'igh*, *liih* 'sand' (*ligig* in Sinkang).

Exceptions (**-S-* is retained as *-h-*): *CaSiq > t<m>ahy 'to sew', *DuSa > rouha, douha 'two'.

Both the phonological innovations, *D, *Z > r, d and loss of intervocalic *-k-* and *-g-* [x], and the suffix *-au* or *-auh* 'future' (Table 3 above) indicate that *The Gospel of St. Matthew* was most likely based on some dialect of Taivuan, rather than Siraya proper. However, that there are exceptions to the loss of intervocalic **-k-* and **-S-* > *-g-*, *-h-* [x] seem to indicate there might be mixture of dialects in *The Gospel of St. Matthew*, as suggested by Adelaar (p.c.), it is "most likely that the *Gospel* text was not the product of one person only: this is clear from the text itself, and ... that there was a committee deciding over the final edition." In fact, it is stated in the introduction to *the Gospel* (p.xiii),

"Hence, too, it follows that the present Translation can be of service to only a few, though populous Villages, such as Soulang, Mattauw, Cinckan (Sinkang), Bacloan, Tavokan, Tevorang, and perhaps also to some of the people in Dorko and Tilocen. These are the places where the work has been carried on for the longest time."

Similarly, it is most likely that *Formulary* (Gravius 1662) was also based on some dialect of Taivuan, whereas *Utrecht Manuscript* was based on a dialect of Siraya, as shown in the following comparison:¹³

¹³ Both Adelaar (1997) and Tsuchida (1998) make a comparison of these three Dutch missionary documents, and Adelaar concludes that *St. Matthew* and *Formulary* are the same language. Adelaar (2006) also notes two main dialects, the "Gospel dialect" and the "Utrecht Manuscript dialect."

Table 6: A Comparison of Utrecht, Matthew, and Formulary

(8)		Utrecht Ms.	St. Matthew	Formulary	
	*D	s	r, d	r, d	
	*Daya	taga-seia	reia	—	'east'
	*DapaN	sapal	rahpal	rahpal	'foot'
	*DuSa	so-soa	dou-rouha	rou-rouha	'two'
	*likuD	ricos	rikour	rikour	'back'
	*lahuD	taga-raos	raour	—	'west'
		sama	rama	dama	'father'
		sa	ra	ra	'but'
		soo, sou	rou	dou, rou	'if, as, when'
		isang	irang	irang	'great, large' ¹⁴
		sasim	rarim	mou-rarim	'down, below'
		pesanach	paeraenaeh	paeraenaeh	'tree'
		massou	—	marou	'corn'
		ka-pousoch-ang	pourough	pourough	'land'
(9)	*-k-	-k-, ø	-k-, ø	ø	
		acoussey	akousi, aousi	aoussi	'have not' ¹⁵
	*(i)aku	iaui, -au	jau, -au	jau	'I'
(10)	*-S-/*-R-	-h-, ø/-g-, -h-	-h-, ø/-h-, ', ø	-h-, ø/h, ', ø	
	*kaSu	cau [kaw]	kow	kow	'thou'
	*DuSa	so-soa	dou-rouha	rou-rouha	'two'
	*CaSiq	t<m>ahy	—	—	'sew'
	*waRi	wagi	wæ'i	wæ'i	'sun, day'
	*wiRi	ougi	u-i	ou-i	'left'
	*baqəRu	vaho	vahæu, va'æu	vahæu, va'æu	'new'
	*kiRim	k<m>igim	k<m>i'im	k<m>i'ym	'seek'
	*kaRaC	k<m>agat	—	—	'bite'
		ligig	li'igh	—	'sand'
		ma-dagoa	—	dæeua	'entirely'
(11)		-a, -al, -ale	-ah, -auh	-a, -ah, -al	'future'

¹⁴ The derived forms of *isang/irang* are *maisasang/maeraerang* in Table 2. It is cognate with *ma-'iDang* 'big' in Puyuma, as suggested by Tsuchida (p.c.).

¹⁵ As based on Adelaar's (p.c.) research, these are two different negatives: *assei*, *assey* or *assi* means 'no(t),' while *akou-si* means 'have-not,' which is derived from *akou-* 'to have' + *(a)ssi* 'not.' Still another negative is *inang* 'will not.'

The examples in (8) show that the language of the *Utrecht Manuscript* is similar to that of Sinkang and Tohkau villages of Siraya, while the language of *St. Matthew* and the *Formulary* is similar to that of Wanli and Matau villages of Taivuan or Lower Tamtsui village of Makatau, as illustrated in (3) above. However, the pronominal forms, first person and second person singular, and the negative in (9) and (10) do not show much difference between the groups. The medial *-k-* is kept or lost in the negative forms in the *Utrecht Manuscript* and *St. Matthew*, but lost in the *Formulary*. All three groups have zero reflexes of *S in *kaSu. On the one hand, the *Utrecht Manuscript* has zero reflex of *S in *DuSa, while *St. Matthew* and the *Formulary* have retained *h* as its reflex. On the other hand, while *-g-*, the reflex of *S or *R, is retained in the lexical forms *wagi* 'sun,' *ougi* 'left,' *k<m>igim* 'seek,' *ligig* 'sand' and *ma-dagoa* 'entirely' in the *Utrecht Manuscript*, it is lost in the forms *wæ'i*, *u-i*, *k<m>i'im*, and *li'igh* in *St. Matthew* and the similar forms in the *Formulary*. That is to say, there is some conflicting evidence. But several examples seem to indicate that the *Utrecht Manuscript* is based on a Siraya dialect, while *St. Matthew* and the *Formulary* are based on a Taivuan dialect. The suffix *-ah* 'future' in (11) also indicates that *St. Matthew* and the *Formulary* are based on a Taivuan dialect. It seems clear that there is some dialect mixture in both *St. Matthew* and the *Formulary*.

4. Relative Chronology and Subgrouping

The sound change PAN *D, *d, *Z > *s* in Siraya, > *r ~ d* in Taivuan and Makatau must have taken place prior to the Dutch occupation of Taiwan (1624-1662), as the phonological difference is manifested in the Dutch missionary documents: *s* in the *Utrecht Manuscript* vs. *r ~ d* in *The Gospel of St. Matthew* and the *Formulary*, as discussed in the preceding section.

The change *-S-, *-R- > [x] (written *g*, *gh*, *h*) or lost may have started in the early seventeenth century because the rule applied to some lexical forms, but not to the others containing the consonant even in the same set of language data as recorded by the Dutch missionaries; see Table 6 above.

PAN *l is retained as *r* rather than *h* or zero (see (1) above) in *St. Matthew*, e.g. *lahuD > *raour* 'west,' *piliq > *peri* 'to choose,' *kalih > *k<m>ari* 'to dig.' That is to say, PAN *l was still retained as *r* in the mid-seventeenth century when *St. Matthew* was translated. In fact, it was still retained in a Wanli text dated 1770, as in *lahuD > *raur* 'west,' and in a Matau text dated 1781, as in *likuD > *mi-likoh* 'to return.' It was not lost until much later when the Japanese started to investigate the languages of the southwestern plains in 1897 or not unlikely a

few decades earlier. So the sound change $*l > h$ or \emptyset was a late innovation limited to Taivuan.

If we take the relative chronology of the sound changes into consideration, then the first split of the Sirayaic group should be Siraya, as shown in Figure 2 below:

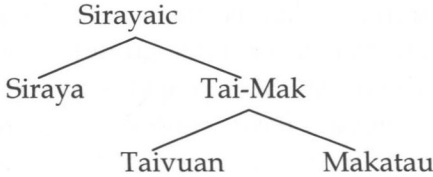


Figure 2: Sirayaic tree diagram based on the chronological order of phonological innovations

However, if we compare the number of shared phonological innovations, then the first split would be Taivuan, as shown in (7) above. Which type of evidence should carry greater weight, the earliest phonological innovation or the number of shared phonological innovations? It seems to me the former should carry more weight in a subgrouping hypothesis.

If the above subgrouping is correct, then the more diverse linguistic area is in the north because both Siraya and Taivuan villages are found in the Tainan Plain to the north. The Sirayaic people may have dispersed to the interior in the east and then later down to Kaoshiung and Pintung prefectures to the south.

Paul Jen-kuei Li
Taipei, Taiwan
March 2010

References

- Adelaar, K. Alexander. 1997. Grammar notes on Siraya, an extinct Formosan language. *Oceanic Linguistics* 36.2:362-397.
- Adelaar, K. Alexander. 1999. Retrieving Siraya phonology: a new spelling for a dead language. *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun & Paul Li, 313-354. Taipei: Institute of Linguistics, Academia Sinica.
- Adelaar, K. Alexander. 2006. Siraya dialogues. *Streams Converging Into an Ocean: Festschrift in Honor of Professor Paul Jen-kuei Li on His 70th Birthday*, ed. by Henry Y. Chang, Lillian M. Huang & Dah-an Ho, 665-686. Taipei: Institute of Linguistics, Academia Sinica.
- Adelaar, K. Alexander. 2007. Siraya, Taiwan's oldest written language. *The Margins of Becoming: Identity and Culture in Taiwan*, ed. by Carsten Storm & Mark Harrison, 19-34. Wiesbaden: Harrassowitz.
- Chen, Bien-horn. 2001. *Notes on Vocabulary of Siraya Version of Matthew*. Association of the Plain Tribes in Tainan. (In Chinese)
- Chen, Bien-horn. 2005. *Notes on Formulary of Christianity in Formosan Siraya Dialect—with translation in English, Chinese and Dutch*. Private Circulation. (In Chinese)
- Gravius, Daniel. 1661. *Het Heylige Euangelium Matthei en Johannis Ofte Hagnau Ka D'llig Matiktik, Ka na sasoulat ti Mattheus, ti Johannes appa. Overgeset inde Formosaansche tale, voor de Inwoonders van Soulang, Mattau, Sinckan, Bacloan, Tavokan, en Tevorang*. Amsterdam: Michiel Hartoch Boekverkoper, inde Oude Hoogstraat, inde Boeck en Papier-winckel.
- Gravius, Daniel. 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, Op den hoeck van de Oude Hooghstraet.
- Li, Paul J. K. 1977. The internal relationships of Rukai. *Bulletin of the Institute of History and Philology Academia Sinica* 48.1:1-92.
- Li, Paul J. K. 1981. Reconstruction of proto-Atayalic phonology. *Bulletin of the Institute of History and Philology Academia Sinica* 52.2:235-301.
- Li, Paul J. K. 1993. New data on three extinct Formosan languages. *Bulletin of the Institute of History and Philology Academia Sinica* 63.2:301-323.
- Li, Paul J. K. 2002. Preliminary interpretations of the 15 recently uncovered Sinkang manuscripts. *Taiwan Historical Research* 9.2:1-68. (In Chinese)
- Li, Paul J. K. 2006. The languages of the plain tribes in southern Taiwan—with