LINGUISTICS OF THE TIBETO-BURMAN AREA
VOLUME 18:1 — SPRING 1995

On the original relationship between Chinese and Kam-Tai

Chen Baoya
Peking University

Does Kam-Tai have any genetic relationship with Chinese? It is a key question both in Sino-Tibetan research and in Asian/Pacific language research. It is also a question involving the methodology of historical linguistics.

Historical linguists of the 19th century used two criteria to determine genetic relationship: one was structural similarity; another was sound correspondences in basic vocabulary. Now, considering language contact and the universal trend of language evolution, more and more linguists do not think of structural similarity as a criterion for determining genetic relationship (Haudricourt 1954, 1961; Matisoff 1970, 1973, 1976; Hu Tan 1980; Dai Qingxia 1980), but sound correspondence is still thought of as the most important criterion. Li Fang Kuei (1937, 1976), Xing Gongwan (邢公畹 1948-1994) and most scholars in China have been sticking to the viewpoint that Chinese and Kam-Tai have a genetic relationship because these two groups share many phonologically and semantically similar words (sound correspondent words).

Having observed the live contact between Dai dialects of Yunnan Province and the Southwest Dialect of Chinese (SDC) for several years, I have found that the SDC loanwords in Dai at the same time and place have strict sound correspondences with original words in SDC. We can take the consonant /tɕh/ of SDC as an example:

Chinese Character	Meaning	SDC	Dai
区	district	/tɕhi^{55}/	/ɕi^{55}/
旗	flag	/tɕhi^{31}/	/ɕi^{42}/
取	to draw (money)	/tɕhi^{53}/	/ɕi^{53}/
撬	sledge	/tɕhiau^{55}/	/ɕau^{55}/
撬	to prize	/tɕhiau^{213}/	/ɕiau^{213}/
翘	to warp	/tɕhiau^{213}/	/ɕiau^{213}/
荞	buckwheat	/tɕhiau^{31}/	/ɕiau^{42}/
请	to invite	/tɕhin^{53}/	/ɕin^{53}/
裙	skirt	/tɕhin^{31}/	/ɕin^{42}/
劝	to advise	/tɕhian^{213}/	/ɕian^{213}/
拳	fist	/tɕhian^{31}/	/ɕian^{42}/
曲	moldy	/tɕhio^{31}/	/so^{31}/
穷	poor	/tɕhiong^{31}/	/song^{31}/

The /tɕh/ of SDC original words corresponds to the /ɕ/ of SDC loanwords in Dai, except for the last two items. These exceptions can be explained. Because /ɕ/ in Dai cannot combine with the diphthongs /io/ and /iong/, /tɕh/ in SDC is pronounced as /s/ in Dai.

Thus the same sound in SDC is pronounced under the same conditions (i.e. in the same phonological environment) as the same sound in loanwords in Dai. This law is very strict. There are two other kinds of phenomena to be explained. Some Dai people pronounce /uəi/, /iəu/, and /ən/ of SDC in a different way:

SDC	*Loan Word in Dai*
/uəi/	/ui/ or /oi/
/iəu/	/iu/ or /eu/
/ən/	/ɯn/ or /ʏn/

The reason for these variants is that in some areas influenced by SDC, Dai is experiencing the convergence of /oi/ and /iu/, /eu/ and /iu/, and /ɯn/ and /ʏn/. In other words, /oi/ and /ui/ in Dai alternate freely; so do /eu/ and /ie/, /ɯn/ and /ʏn/. Therefore, SDC words with /uəi/, /iəu/ and /ən/ each have two kinds of pronunciation in Dai.

As soon as this kind of Dai sound change has run its course, /uəi/, /iəu/ and /ən/ in original SDC words will keep respectively strict correspondents to loanwords in Dai. This still means that the same sounds of original SDC words, under the same conditions, are pronounced as the same sound in loanwords.

Since original words and loanwords have strict sound correspondences, we cannot distinguish genetic relationship from contact relationship only by sound correspondence, nor can we tell cognate words from loanwords in this way.

Some scholars think that basic words (especially 'kernel words' or 'core vocabulary') cannot be borrowed (Li Fang Kuei 1976), so that basic words or kernel words can be used to distinguish genetic relationship from contact relationship. This is unfortunately not true. The 200 words given by M. Swadesh (1952) have been considered as standard kernel words by most linguists, but I have observed, in contact between SDC and Dai in Dehong in Yunnan, that some SDC kernel words have entered Dai, such as /xa:i^{53}/ 'sea', /xun^{53}/ 'tie', /phiau^{55}/ 'float', /te^{55}/ 'father'. The Dai living by Jinsha River in Yunnan have borrowed more SDC kernel words: /tsha^{11}/ 'rub', /lan^{33}/ 'rotten', /ɕiang^{55}/ 'think', /xai^{55}/ 'sea', /tshə^{55}/ 'pull', /sau^{55}/ 'few', /phau^{55}/ 'swell', /tshan^{55}/ 'dig', /tsuan^{33}/ 'turn', /khan^{55}/ 'cut'. All these borrowed words show strict sound correspondences with the original words.

In recent years, Yan Xuequn (1979) and Dong Weiguang (1984) have proposed "correspondence of word groups", while Xing Gongwan (1983)

has proposed "deep structure correspondence". They attempt to use these methods to demonstrate the genetic relationship between Chinese and Kam-Tai. However, I have found that both of these types of correspondences also appear in the borrowed words (Chen Baoya 1994:1.3).

Therefore, there is no reason to set a vocabulary boundary beyond which no words can be borrowed. We maintain that even items of core vocabulary showing regular sound correspondences are not enough to distinguish genetic relationship from contact relationship. If we find two groups of sound correspondent words at different times between two languages, such as those between Kam-Tai and Old Chinese and between Kam-Tai and SDC, we can say the sound correspondent words at the later time are borrowed words because they have different kinds of correspondences from those between Kam-Tai and Old Chinese, but we still have no methods strong enough to tell if the sound correspondent words in earlier time between Kam-Tai and Old Chinese are cognate words or borrowed words. Li Fang Kuei (1976) found more than 100 sound correspondent kernel words between Old Chinese and Tai, 29 of them belonging to Swadesh's 200 word list. Li was more inclined to consider them evidence of the genetic relationship between Chinese and Tai. Since items of core vocabulary can also be borrowed, we think Li deduced too strong a conclusion from his data. Most Chinese scholars explain the relationship between Chinese and Kam-Tai as Li Fang Kuei did.

Now we can conclude that contact is potenetially unlimited. "Unlimited" means that any language level, including kernel structure and kernel vocabulary, can be influenced in contact. Therefore regular sound correspondences are only a necessary condition for genetic relationship, not a sufficient one.

But this does not mean that we cannot distinguish genetic relationship from contact relationship at all. According to my personal observations of contemporary contact phenomena, both language contact and language division may be ranked on a scale, and the two kinds of scales are opposite. In language contact, the more nuclear sound correspondent words are fewer than the less nuclear ones. In language division, on the other hand, nuclear sound correspondent words are more numerous than less nuclear ones.

To avoid personal bias in selecting items of core vocabulary, I have divided Swadesh's 200 kernel words into two groups to see the different semantic patterning in situations of language division versus language contact. In 1952, Swadesh listed 200 words which he thought to be the most stable words in human languages. But even these 200 words could be borrowed from one language to another. In 1955, Swadesh narrowed his list down to 100 words. He thought that these 100 words were the most stable and could not be borrowed. As we discussed above, even these 100 words

can also be borrowed as a result of Chinese and Tai contact. I have divided Swadesh's 200 kernel words into two ranks:

1st 100 words: Swadesh's 100 kernel words (1955)
2nd 100 words: the remaining items on Swadesh's list of 200 words (1952)

Having subtracted Swadesh's 100 words from his 200-word list, we are left with 107 words. I have deleted seven of them, for in Asian languages they are either partially repetitive with others in meaning or inconvenient to be compared. These are: *at* (partial overlap with *in*), *when* (partial overlap with *what*), *wipe* (partial overlap with *rub*), *with* (partial overlap with *and*), *ye* (partial repetition with *you*), and *some, other* (inconvenient to compare).

I have found that in genetic relationship, the cognates among the first 100 words are more numerous than those in the second 100 words, while in a contact relationship, the loanwords among the first 100 words are fewer than those in the second 100 words.

Let us first look at the distribution of kernel cognate words in Dai (=Tai) dialects of Yunnan Province:

	DD *Dehong*	**DY** *Yun River*	**DL** *Lincan*	**DJ** *Jinsha River*
DX *Xishuangbanna*	88/71	91/68	85/71	84/66
DD *Dehong*		92/72	94/85	91/69
DY *Yun River*			91/68	88/68
DL *Lincan*				88/68

On the left of the slashes are the numbers of cognate words in the first 100 words; on the right of the slashes are those in the second 100 words. It is evident that the cognates among the first 100 words are more numerous than those in the second 100 words.

Now let's see how core cognates in some European languages whose genetic relationship has been proven are distributed in the first 100 words and the second 100 words. In 1952 and 1955, Swadesh gave out the following figures:

	Cognate words	
	in 200 word list	*in 100 word list*
Old English/Modern English	77%	86%
Old German/Modern German	84%	89%
Old Swedish/Modern Swedish	85%	94%
Latin/Modern Romanian	56%	71%
Latin/French	62%	74%
Old Greek/Modern Greek	69%	71%

According to the formula $x_2=2x-x_1$ (where 2x is the number of cognates in the 200 word list, x_1 is the number of cognates in the first 100 word list, and x_2 is the number of words in the second 100 word list), we can figure out cognate percentages as follows:

	Cognate words	
	in first 100 words	*in second 100 words*
Old English/Modern English	86%	68%
Old German/Modern German	89%	79%
Old Swedish/Modern Swedish	94%	76%
Latin/Modern Romanian	71%	41%
Latin/Modern French	74%	50%
Old Greek/Modern Greek	71%	67%

There is no doubt that all these pairs of languages have a genetic relationship. All the figures show that, in each pair, cognates among the first 100 words are more numerous than those in the second 100 words.

There is also no doubt that English belongs to the Germanic group. The following figures show the distribution of cognate words between English and some other languages of the Germanic group in the first 100 words and the second 100 words (judgments of cognacy are according to Barnhart 1988).

	Cognate words	
	in first 100 words	*in second 100 words*
Old English/English	87%	71%
German/English	64%	46%
Old Frisian/English	72%	51%
Dutch/English	71%	54%
Old Icelandic/English	76%	57%
Gothic/English	58%	39%

In each case, the cognates among the first 100 words are more numerous than those in the second 100 words.

Let's turn to Chinese. It is obvious that the Chinese dialects have a genetic relationship to each other. Compare the distribution of their kernel cognate words:

	Hakka	*Southern Min*	*Xiang*	*Wu*	*Gan*	*Yue*
Mandarin	75/56	61/56	82/71	82/71	76/73	83/71
Hakka		69/51	72/52	71/59	66/56	70/55
Southern Min			65/51	62/52	62/49	60/49
Xiang				87/64	81/67	78/68
Wu					78/72	76/72
Gan						71/69

On the left of the slashes are the percentages for the first 100 words; on the right, those for the second 100 words. These figures also show that the cognates in the first 100 words are more numerous than those in the second 100 words.

By any criterion, the Indo-European languages and Chinese dialects are among the most important language groups in the world; it is striking that in both groups, there are more genetic cognates in the first 100 words than in the second 100 words.

Now let's turn to contact relationship. In the contact between the Southwest Dialect of Chinese and different dialects of Tai, all the dialects of Tai have borrowed kernel words from the Southwest Dialect of Chinese in different ways and to varying degrees. I have collected those in Dehong (DD), Jinsha (DJ) and Yun River (DY):

Loans from Southwest Dialect of Chinese into Dai in Dehong

Meaning	*SDC originals*	*Loans in DD*	*Collocations*
First 100 words:			
hair	毛 /mau^{31}/	/mau^{42}/	/mau^{42}ɕian^{213}/ 'knitting wool' /mau^{42}pi^{31}/ 'writing brush'
liver	肝 /kan^{55}/	/kan^{55}/	/kan^{55}ian^{55}/ 'hepatitis'
Second 100 words:			
father	爹 /tie^{55}/	/te^{55}/	
float	漂 /phiau^{55}/	/phiau^{55}/	
dull	憨 /xan^{55}/	/xan^{55}/	
turn	转 /tsuan^{213}/	/tsuan^{213}/	
ice	冰 /pin^{55}/	/pin^{55}/	
sea	海 /xai^{213}/	/xai^{213}/	
tie	捆 /khun^{53}/	/xun^{53}/	
squeeze	压 /jia^{31}/	/jia^{31}/	
if	如果 /zu^{31}ko^{53}/	/zu^{31}ko^{53}/	
because	因为 /jin^{55}uei^{213}/	/jin^{55}vui^{213}/	

Loans from Southwest Dialect of Chinese into Dai of Jinsha

Meaning	*SDC original*	*Loans in DJ*
First 100 words:		
(no borrowed words)		
Second 100 words:		
rub	擦 /tsha^{31}/	/tsha^{11}/
rotten	烂 /lan^{213}/	/lan^{33}/
think	想 /ɕia^{53}/	/ɕia^{55}/
sea	海 /xai^{53}/	/xai^{55}/
split	扯 /tshə^{53}/	/tshə^{55}/
few	少 /ʂau^{53}/	/ʂau^{55}/
swell	肿 /phau^{55}/	/phau^{55}/
turn	转 /tshuan^{213}/	/tshuan^{33}/
cut	砍 /kan^{53}/	/kan^{55}/

Loans from Southwest Dialect of Chinese into Dai of Yun River

Meaning	*SDC originals*	*Loans in DY*	*Collocations*
First 100 words:			
oil; grease	油 /jəu^{31}/	/jəu^{33}/	/pan^{33}jəu^{33}/ 'pork fat'
kill	杀 /ʂa^{31}/	/ʂa^{31}/	
water	水 /ʂw^{31}/	/ʂwəi^{53}/	/pən^{53}ʂwəi^{53}/ 'water in ditch'
Second 100 words:			
grass	草 /tshau^{53}/	/tsha:u^{53}	/tsha:u^{53}xa:i^{31}/ 'straw sandals'
think	想 /ɕiang^{53}/	/ɕa:ng^{53}/	
sea	海 /xai^{53}/	/xa:i^{53}/	/xa:i^{53}tsɿ^{53}/ 'lake'
and	和 /xo^{31}/	/xo^{31}/	
old	老 /lau^{53}/	/la:u^{53}/	/la:u^{53}pu^{24}thau^{11}/ 'old man' /la:u^{53}ja^{53}/ 'old woman'
squeeze	压 /ja^{31}/	/ja^{33}/	
father	爹 /tie^{55}/	/ʔa^{33}tje^{33}/	
dull	笨 /pən^{213}/	/pən^{11}/	

The percentages reflected by the above data are:

	DD *Dehong*	**DJ** *Jinsha River*	**DY** *Yun River*
SDC *Southwest Dialect of Chinese*	2/10	0/9	3/8

All the data show that the cognates among the first 100 words are more numerous than those in the second 100 words, and that conversely loanwords among the first 100 words are more numerous than those in the second 100 words.

So far we can conclude that the distribution of kernel loanwords is opposite to that of kernel cognate words.

We have seen that both cognate words and loanwords can show regular sound correspondences, so that this criterion is insufficient by itself to distinguish cognate words from loanwords, or genetic relationship from contact relationship. However, now that we see that the semantic distribution of kernel cognate words is opposite to that of kernel loanwords,

we can distinguish genetic relationship from contact relationship by this opposition. The method is, when confronted with two languages whose genetic or contact relationship needs to be determined, if the sound correspondent words in the first 100 words are more numerous than those in the second 100 words, we can say the two languages have a genetic relationship. Conversely, if the sound correspondent words in the first 100 words are fewer than those in the second 100 words, we can say the two languages have a contact relationship.

Most linguists believe that there is a Proto-Kam-Tai, including a Tai (Zhuang-Dai) branch, a Kam-Sui (Dong-Shui) branch and a Li (Hlai) branch. I have selected eleven languages from Kam-Tai to observe the distribution of kernel sound correspondent words. These languages are:

Tai Group:	ZW	Zhuang in Wuming (Guangxi)
	ZL	Zhuang in Longzhou (Guangxi)
	BY	Buyi (Guizhou)
	DX	Dai in Xishuangbanna (Yunnan)
	DD	Dai in Dehong (Yunnan)
Kam Group:	DR	Dong in Rongjiang (Guizhou)
	MLL	Mulao in Luocheng (Guangxi)
	SS	Sui in Sandu (Guizhou)
	MLH	Maolan in Huangjiang (Guangxi)
Li (Hlai) Group:	LB	Li in Baoding (Hainan)
	LT	Li in Tongshi (Hainan)

Now compare the following percentages of sound correspondent words in the first 100 words and the second 100 words:

First 100 words:

	ZL	BY	DX	DD	DR	MLL	SS	MLH	LB	LT
ZW	86	90	78	76	61	56	57	56	49	46
ZL		78	80	72	54	52	50	46	48	46
BY			72	72	56	52	56	54	52	50
DX				88	52	51	53	48	48	49
DD					48	48	50	47	51	48
DR						74	80	79	40	39
MLL							76	73	38	37
SS								79	37	38
MLH									37	38
LB										90

Second 100 words:

	ZL	BY	DX	DD	DR	MLL	SS	MLH	LB	LT
ZW	69	81	53	54	46	48	54	55	27	32
ZL		61	55	58	38	40	41	41	25	24
BY			53	51	46	47	52	52	25	27
DX				71	36	34	44	37	26	29
DD					34	33	40	37	31	30
DR						56	59	57	18	21
MLL							50	59	18	18
SS								62	21	24
MLH									19	22
LB										90

The figures show that in each pair of 11 languages selected from Kam-Tai, the correspondent words in the first 100 words are more numerous than those in the second 100 words. We should conclude that each pair of 11 languages has a genetic relationship, and there was indeed a Proto-Kam-Tai.

Let us turn to the most controversial issue in Sino-Tibetan research; that is, what kind of relationship exists between Chinese and Kam-Tai. I have compared the kernel words of Old Chinese (OC) with those of eleven languages of Kam-Tai and have obtained the following result: (The single numbers refer to classes of tones; OC reconstructions are according to Wang Li).

Comparison of Old Chinese kernel words with those of Kam-Tai languages

First 100 words:

Chinese	Meaning	OC	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
角	horn	kɔk^{7}		ko:k^{7}							
路	road	lu^{8}	lo^{6}	lo^{6}							
绿	green	liwok^{8}	lok^{8}								
尔	you	ɾie^{4}		ni^{6}							
爪	claw	tʃau^{3}						cɐu^{3}	tsa:u^{3}		
白	white	bɐk^{8}	pi:k^{8}	phə:k^{7}			phək^{9}	pa:k^{10}	pa:k^{8}	pa:k^{8}	pok^{8}
皓	white	ʏau^{4}	ha:u^{1}	kha:u^{2}	ʏa:u^{1}	xa:u^{1}	xa:u^{1}				
地	earth	di^{6}	tei^{6}	ti^{6}			ti^{6}	ti^{6}	ti^{6}	ti^{6}	ti^{6}
二	two	ɾi^{6}	ngei^{6}	ji^{6}	ni^{6}			ni^{6}	ni^{6}	ni^{6}	ni^{6}
皮	skin	bie^{2}						pi^{2}		pi^{2}	pi^{2}
热	hot	ɾiɛt^{8}							net^{8}	sut^{7}	
沙	sand	ʃa^{1}	sa^{1}								
听	hear	thieng^{1}	ting^{5}	ting^{6}				thiŋg^{5}	theng^{5}		
我	I	nga^{4}		ngo^{6}							
心	heart	siem^{1}	sim^{1}	lim^{1}				səm^{1}	təm^{1}	cum^{1}	sam^{1}
一	one	iet^{7}	it^{7}	it^{7}	it^{7}	et^{7}	et^{7}	et^{7}	ʔjət^{7}		ʔjit^{7}
油	grease	jiəu^{2}	jou^{2}	ju^{2}	jiu^{2}			ju^{2}	jəu^{2}		ju^{2}
名	name	miɛng^{2}	ming^{2}	ming^{2}					mɛ:ng^{2}		
字	name	dziə^{6}	ɕo^{6}		tso^{6}	tsɯ^{6}	tsɯ^{6}				
那	that	nɑ^{6}				na^{3}					
人	man	ɾien^{2}					njən^{2}	cəɲ^{1}	zen^{1}	zən^{1}	
乳	breasts	ɾio^{4}					no^{4}				
咬	bite	ngau^{4}							nga:u^{6}		
觉	sleep	keəuk^{5}								kau^{2}	kau^{2}

Second 100 words:

Chinese	Meaning	OC	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
浮	float	biəu^{2}	fou^{2}	fu^{2}	vu^{2}	fu^{2}	fu^{2}		fu^{2}		
父	father	biu^{6}	po^{6}	pa^{5}	po^{6}	pɔ^{6}	po^{6}	pu^{4}	pu^{4}	pu^{4}	
旧	old	giəu^{6}	kau^{5}	kau^{5}	kau^{5}	kau^{5}	kau^{5}	a:u^{5}	ko^{5}	qa:u^{5}	ka:u^{5}
股	leg	ku^{3}	ka^{1}	kha^{1}	ka^{1}	xa^{1}	xa^{1}	pa^{1}		pa^{1}	pja^{1}
平	smooth	bieng^{2}	ping^{2}	phi:ng^{2}	ping^{2}	peng^{2}	peng^{2}	pjing^{2}	peng^{2}	pjeng^{2}	peng^{2}
吸(吮)	suck	xiəp^{7}	ɕup^{7}		ɕup^{7}						
挖	dig	wæt^{7}	vat^{7}	vat^{7}					wa:t^{7}		ʔwa:t^{7}
掘	dig	giwət^{8}	kut^{8}			xut^{8}	xot^{8}				
短	short	tuan^{3}	tin^{3}	tin^{3}	tin^{2}		sɛn^{3}	thən^{3}	hʏən^{3}	ndjən^{3}	din^{4}
棍	stick	kuən^{5}								kwən^{5}	
海	sea	xɒi^{3}	ha:i^{3}	ha:i^{3}	ʏa:i^{3}			həi^{3}	hai^{3}	ha:i^{3}	hai^{3}
河	river	ʏɑ^{2}					xe^{2}				
华(花)	flower	xwa^{1}	va^{1}		va^{1}			wa^{1}	hwa^{1}	wa^{1}	
坏	bad	ʏwɐi^{6}	va:i^{6}	va:i^{6}	va:i^{6}			wa:i^{6}	hwa:i^{5}		wa:i^{5}
尖	sharp	tsiɛm^{1}	li:m^{1}	ɬi:m^{3}	som^{1}	lɛm^{1}	lɛm^{1}				sam^{1}
结(冰)	freeze	kiet^{7}	ki:t^{7}		tɕi:t^{7}			tjit^{9}	cet^{7}	tjet^{7}	cet^{7}
老	old	lɑu^{4}	la:u^{4}		la:u^{4}			la:u^{4}	lo^{4}	la:u^{4}	la:u^{4}
年	year	nien^{2}						njin^{2}	njen^{2}		
少	few	ɕiɛu^{3}	siu^{3}		siu^{3}				sjèu^{3}	sjeu^{3}	ɕeu^{3}
索	rope	sɑk^{7}/sɐk^{7}	ɕa:k^{8}	tɕɔ:k^{8}	tsɔ:^{8}	tsək^{8}			luk^{7}	la:k^{7}	za:k^{7}
四	four	si^{5}	sei^{5}	ɬi^{5}	si^{5}	si^{5}	si^{5}	si^{5}	ti^{5}	ɕi^{5}	si^{5}
他	he	thɑ^{1}	te^{1}		te^{1}						
推	push	thuɒi^{1}							thoi^{1}		
吹(风)	blow	tɕhiwe^{1}							tshui^{1}		
五	five	ngu^{4}	ha^{3}; ngu^{4}	ha^{3}	ʏa^{3}	ha^{3}	ha^{3}	ngo^{4}	ngɔ^{4}	ngo^{4}	ngɔ^{4}
窄	narrow	tʃɐk^{7}		tɕek^{7}							
夫(丈夫)	husband	piu^{1}		po^{6}		pho^{1}	pho^{1}				
擦	rub					tset^{8}	tset^{8}	sa:t^{9}		dja:t^{7}	

唱	sing	tɕhiang^{5}	ɕi:ng^{5}	ɕɔ:ng^{5}				tshja:ng^{5}	ɕi:ng^{5}	
虻(虫)	worm	mɐng^{2}		meng^{2}	mɛng^{3}	mɛng^{3}				
生(活的)	alive	ʃɐng^{1}						sɛ:ng^{1}		
烂	rotten	lan^{6}			la:n^{6}				la:n^{6}	
想	think	siang^{3}	si:ng^{3}	ɬa:ng^{3}			ɕa:ng^{5}	tja:ng^{5}		sə:ng^{3}
雪	snow	siwet^{7}		li:t^{7}						

The percentages of correspondent words with Chinese are:

	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
First 100 words	13	13	5	4	7	11	10	6	7
Second 100 words	22	19	18	15	15	15	22	16	15

It is obvious that the sound correspondent words among the first 100 words are fewer than those in the second 100 words. We should confess that the intimate relationship between Chinese and Kam-Tai is a contact relationship.

Some scholars suppose that Kam-Tai and Austronesian have a genetic relationship (see Benedict 1942: 1975). I have selected Malay (M) and Indonesian (I) from Austronesian to compare with Kam-Tai. Observe the following data:

Comparison of Indonesian and Malay kernal words with those of Kam-Tai languages

First 100 words:

Meaning	I	M	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
black	hitam	hitam	dam^{3}	dam^{1}		dam^{1}	dam^{6}	nam^{1}	nam^{1}	ʔnam^{1}	nam^{1}
fire	api	api	fei^{2}	fai^{2}	vi^{2}	fai^{2}	fai^{2}	pui^{1}	fi^{1}	vi^{1}	vi^{1}
green	hidau	hijau	heu^{1}	kheu^{1}		xeu^{1}	xeu^{1}	su^{1}	həu^{1}	ɕu^{1}	ju^{1}
water	zamzam		ɣam^{4}	nam^{4}	zam^{4}	nam^{4}	lam^{4}	nam^{4}	nam^{4}	nam^{3}	nam^{3}
die	mati	mati	ta:i^{1}	ha:i^{1}	ta:i^{1}	ta:i^{1}	ta:i^{6}	təi^{1}	tai^{1}	tai^{1}	tai^{1}
I	aku	aku	kou^{1}	kau^{1}	ku^{1}	ku^{1}	kau^{6}	ja:u^{2}	həi^{2}	ju^{2}	hou^{2}
eye	mata	mata	ta^{1}	ma:k^{7}ha^{1}	ta^{1}	ta^{1}	ta^{6}	ta^{1}	la^{1}	nda^{1}	nda^{1}
moon	bulan	bulan	ɣo:ng^{8} duɯ:n^{1}		zuang^{6} duɯ:n^{1}	dən^{1}	lən^{6}	nja:n^{1}	kʏa:ng^{2} njen^{2}	njen^{2}	ni^{4}njen^{2}
this	ini	ini	nei^{4}	nai^{3}	ni^{4}	ni^{4}	lai^{4}	na:i^{6}	na:i^{6}	na:i^{6}	na:i^{6}
nose	hidung	hidong	dang^{1}	mə:k^{7} dang^{1}	dang^{1}	hu^{2}dang^{1}	hu^{2}lang^{6}	nang^{1}	kə^{5} nang^{1}	ʔnang^{1}	nang^{1}
eat	makan	makan	kɯn^{1}	kin^{1}	kɯn^{1}	kin^{1}	kin^{6}	tja:n^{1}	tsa:n^{1}		
you	kamu	-mu	mɯng^{2}	maɯ^{2}	mɯng^{2}	mɯng^{2}	maɯ^{2}				
new	baru	baharu	mo^{5}	maɯ^{5}	mo^{5}	mai^{5}	maɯ^{5}	ɯəi^{5}	mai^{5} tui^{2}	mai^{5} tui^{2}	mai^{5}
stone	batu	batu									
know	tahu		ɣo^{4}	ɬu^{4}na^{3}	zo^{4}	hu^{4}	hu^{4}	wo^{4}	ɣɔ^{4}ceu^{3}		wo^{3}
come	mari			ma^{2}	ma^{1}	ma^{2}	ma^{2}	ma^{1}			

Second 100 words:

Meaning	I	M	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
child	anak	anak	luɪk^{8}nje^{2}	luk^{8}dik^{7}	luɪʔ^{8}sai^{5}	luk^{8}ɔn^{5}	luk^{8}	la:k^{10}un^{3}	la:k^{8}te^{5}	la:k^{8}ti^{3}	la:k^{8}ce^{3}
fall	dzatuh	jatoh	tok^{7}	tuk^{7}	toʔ^{7}	tok^{7}	tok^{9}	tok^{7}		tok^{7}	tɔk^{7}
leg	paha	paha	ka^{1}	kha^{1}	ka^{1}	xa^{1}	xa^{1}	pa^{1}		pa^{1}	pja^{1}
lake	kolam	kolam	tam^{2}	thum^{1}	tam^{1}			tam^{1}	lam^{1}	ndam^{1}	ndam^{1}
salt	garam	garam	kju^{1}	kuɪ^{1}	tɕu^{1}	kə^{1}	kə^{8}	ko^{1}	cwa^{1}	kwa^{1}	kwo^{1}
wife	bini	bini	me^{8}ja^{8}	me^{8}	ja^{8}	me^{2}	me^{2}	ma:i^{4}	ma:i^{4}	ni^{4}ja^{4}	lja^{3}

The percentages of Kam-Tai/Indonesian correspondent words are:

	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
First 100 words	14	14	13	15	15	14	14	11	13
Second 100 words	6	6	6	5	5	6	4	6	6

The percentages of Kam-Tai/Malay comparabilia are:

	ZW	ZL	BY	DX	DD	DR	MLL	SS	MLH
First 100 words	12	11	10	15	12	12	11	10	11
Second 100 words	6	6	5	5	5	6	4	6	5

According to these data, we may say that Austronesian and Kam-Tai have a genetic relationship.

Now we can conclude in the light of these scales of kernel words that there are two important family trees in the Asia/Pacific region: one is Sino-Tibetan; the other is "Yue", including Austronesian and Kam-Tai. These two linguistic stocks have been in very close contact, so there is much similarity between them. Most Chinese scholars think that Kam-Tai belongs to Sino-Tibetan because of structural similarity and sound correspondences. Most European and American scholars classify Chinese and Kam-Tai into different families because of sound correspondences between Kam-Tai and Austronesian. We have seen that structural similarity is not enough by itself to distinguish a genetic relationship from a contact relationship, nor are regular sound correspondences. To distinguish the two kinds of relationship, we must refer to the semantic patterning of resemblant items of core vocabulary.

REFERENCES

Bao Sijing 包思井. 1952. *Practical Chinese-Indonesian Dictionary* 实用华印辞典. Yin-Ma Bookstore 印马书店.

Barnhart, Robert K. 1988. *The Barnhart Dictionary of Etymology*. New York: H. W. Wilson Co.

Benedict, Paul K. 1942. "Thai, Kadai and Indonesian: a new alignment in Southeastern Asia." *American Anthropologist* 44.

_____. 1972. *Sino-Tibetan: a Conspectus.* Contributing editor, James A. Matisoff. Cambridge: Cambridge University Press.

_____. 1975. *Austro-Thai Language and Culture, with a Glossary of Roots.* New Haven: Human Relations Area Files Press.

Central Institute of Nationalities, Research Section Five, Minority Languages Division 中央民族学院少数民族语言研究所第五研究室. 1985. *Collected Word Lists of Kam-Tai Languages* 壮侗语族语言词汇集.

Chang Kun 张昆. 1977. "Current trends in the study of non-Chinese languages in China" 中国境内非汉语研究的方向. *Collected Works in Chinese Linguistics* 中国语言学论集. Taiwan: Youshi Wenhua Shiye Gongsi 幼狮文化事业公司.

Chen Menghe 陈蒙鹤 et al. 1961. *Chinese-Malay Dictionary* 华马大辞典. Singapore: Shanghai Books Limited 上海书局有限公司.

Chen Baoya 陈保亚. 1993. "The language alliance between Chinese and Tibetan" 汉藏语言联盟说. *Culture, History and Folklore* 文化、历史、民俗. Yunnan University Press 云南大学出版社.

_____. 1993. *On Language Culture* 语言文化论. Yunnan University Press 云南大学出版社.

_____. 1993. "*Qiangxia*: On the linguistic and cultural 'alliance' within Austro-Thai ('Ao-Yue')" 羌夏－澳越语言文化联盟论. *Bulletin of the Yunnan Institute of Nationalities* 云南民族学院学报 1993.3.

_____. 1994. *Language Contact and Language Alliance* 论语言接触与语言联盟. *Beijing: Peking University dissertation.*

Dai Qingxia 戴庆厦. 1980. "A study of tense and lax vowels in Tibeto-Burman" 藏缅语族松紧元音研究. *Selected Papers of the Central Institute of Nationalities* 中央民族学院学术论文选.

Dong Weiguang 董为光、Cao Guangqu 曹广衢、Yan Xuequn 严学　. 1984. "The genetic relationship between Chinese and Kam-Tai." *Computational Analyses of Asian and African Languages* 22:105-121. Tokyo.

Hall, R. A. 1958. "Creolized languages and genetic relationship." *Word* 14:367-373.

Hancock, Ian F. 1980. "Lexical expansion in creole language." *Theoretical Orientations in Creole Studies*, ed. by Albert Valdman and Arnold Highfield, 63-88. New York: Academic Press.

Haudricourt, A-G. 1954. "De l'origine des tons en vietnamien." *Journal Asiatique* 242:69-82.

_____. 1961. "Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient." *Bulletin de la Société de Linguistique de Paris* 56.1:163-18.

Haugen, Einar. 1950. "The analysis of linguistic borrowing." *Language* 26:210-231.

Hu Tan 胡坦、Qu Aitang 瞿霭堂、Lin Lianhe 林联合. 1980. "A study of tone in Tibetan (Lhasa Dialect)" 藏语（拉萨话）声调研究. *Minzu Yuwen* 民族语文 1980.1.

Hymes, Dell. 1955. "Positional analysis of categories: a frame for reconstruction." *Word* 11:10-23.

_____. 1971 *Pidginization and Creolization of Languages.* Cambridge: Cambridge University Press.

Li, Fang Kuei 李方桂. 1937. *Language and Dialects of China, Chinese Yearbook.* Shanghai: Commercial Press.

_____. 1940. "The Longzhou Dialect of Zhuang" 龙州土语. *Qian Shiyusuo Dankan Jiazhong zhi Shiliu* 前史语所单刊甲种之十六.

_____. 1943. "Notes on Mo Dialect" 莫话记略. *Qian Shiyusuo Dankan Jiazhong zhi Ershi* 前史语所单刊甲种之二十.

_____. 1964. "Tai and Kam-Sui" 台语与侗水语. *Lingua* 14.1.

_____. 1965. "The Tai and Kam-Sui languages." *Indo-Pacific Linguistic Studies*, Vol. I, ed. by G. B. Milner and Eugénie J. A. Henderson, 148-179. Amsterdam: North-Holland Publishing Co.

_____. 1976. "Sino-Tai." *Computational Analyses of Asian and African Languages* 3:39-48.

_____. 1977. *A Handbook of Comparative Tai.* Honolulu: University of Hawaii Press.

_____. 1980. *A Study of Old Chinese Phonology* 上古音研究. Beijing: Shangwu Yinshuguan 商务印书馆.

Luo Meizhen 罗美珍. 1983. "The problem of the genetic affiliation of Tai" 试论台语的系属问题. *Minzu Yuwen* 民族语文 1983.2.

_____. 1992. "Identifying Chinese-Kam cognates" 汉傣同源词辨. *Language Study and Usage* 语言研究与应用. Beijing: Shangwu Yinshuguan 商务印书馆.

_____. 1993. "On dialects: A joint discussion on the differentiation of Dai dialects" 论方言——兼谈傣语方言的划分. *Minzu Yuwen* 民族语文 1993.3.

Matisoff, James A. 1970. Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case study. *Journal of the American Oriental Society* 9.1:13-44.

_____. 1972. *The Loloish Tonal Split Revisited.* Monograph #7. Berkeley: Center for South and Southeast Asian Studies.

_____. 1973. "Notes on Fang-kuei Li's 'Languages and dialects of China.'" *Journal of Chinese Linguistics* 1.3:471-474.

_____. 1973. "Tonogenesis in Southeast Asia." *Consonant Types and Tone,* ed. by Larry M. Hyman, 71-96. (Southern California Occasional Papers in Linguistics, 1.) Los Angeles: University of California.

_____. 1974. "The tones of Jinghpaw and Lolo-Burmese: common origin vs. independent development." *Acta Linguistica Hafniensia* 15.2:153-212.

_____. 1975. "Benedict's Sino Tibetan: a rejection of Miller's conspectus inspection." *LTBA* 2.1:155-172.

_____. 1976. "Austro-Thai and Sino-Tibetan: an examination of body-part contact relationships." *Genetic Relationship, Diffusion, and Typological Similarities of East and Southeast Asian Languages,* ed. by Mantaro J. Hashimoto, 256-289. Tokyo: Japan Society for the Promotion of Science.

_____. 1991. "Jiburish revisited: tonal splits and heterogenesis in Burmo-Naxi-Lolo checked syllables." *Acta Orientalia* 52:92-114.

_____. 1991. "Sino-Tibetan linguistics: present state and future prospects." *Annual Review of Anthropology* 20:469-504.

Mühlhäusler, P. 1980. "Structural expansion and the process of creolization." *Theoretical Orientations in Creole Studies*, ed. by Albert Valdman and Arnold Highfield, 19-55. New York: Academic Press.

Ni Dabai 倪大白. 1988. "Kam-Tai and Austronesian in China" 中国的壮侗语与南岛语. *Bulletin of the Central Institute of Nationalities* 中央民族学院学报 1988.3.

_____. 1990. *Outline of Kam-Tai languages* 侗台语概论. Central Institute of Nationalities Press 中央民族学院出版社.

Peking University Chinese Department, Linguistics Teaching and Research Section 北京大学中文系语言学教研室. 1964. *Word List of Chinese Dialects* 汉语方言词汇. Wenzi Gaige Chubanshe 文字改革出版社.

Ratanakul, Suriya. 1986. *Southeast Asian Linguistic Studies*. Bangkok: Mahidol University.

Sagart, Laurent. 1990. "Chinese and Austronesian are genetically related!" Paper presented at the 23rd ICSTLL, Arlington, Texas.

Swadesh, Morris. 1952. "Lexicostatistic dating of prehistoric contacts, with special reference to North American Indians and Eskimos." *Proceedings of the American Philosophical Society* 96:452-463.

_____. 1954. "Time depths of American linguistic groupings." *American Anthropologist* 56.

Thomason, Sarah G. and Terrence S. Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley and Los Angeles: University of California Press.

Valdman, Albert and Arnold Highfield. 1980. *Theoretical Orientations in Creole Studies*. New York: Academic Press.

Wang Futang 王福堂. 1995. *Word List of Chinese Dialects* 汉语方音词汇. Language Press 语文出版社.

Wang Jun 王均 et al. 1984. *A Brief Introduction to the Kam-Tai Languages* 壮侗语族语言简志. Nationalities Press 民族出版社.

Weinreich, Uriel. 1953. *Languages in Contact.* The Hague: Mouton.

Wen Xitao 温锡涛. 1951. *Chinese-Indonesian Dictionary* 中印大辞典. Nanbei Wenhua Chubanshe 南北文化出版社.

Winter, Werner. 1973. "Areal linguistics: some general considerations." *Current Trends in Linguistics, Vol. 11: Diachronic, Areal, and Typological Linguistics*, ed. by Thomas A. Sebeok, 135-147. The Hague: Mouton.

Xing Gongwan 邢公畹. 1948. "A preliminary explanation of Chinese *zi, er* and the Tai particle *luk*" 汉语"子"、"儿"和台语语助词 luk 试释. *Guowen Monthly* 国文月刊 68.

_____. 1949. "A comparative study of Chinese and Tai morphology" 汉台语构词法的一个比较研究. *Guowen Monthly* 国文月刊 77.

_____. 1962. "On the importance of tonal categories for Chinese-Tai comparative studies" 论调类在汉台语比较研究上的重要性. *Zhongguo Yuwen* 中国语文 111.

_____. 1973. "The particles *le* and *zhe* in Modern Chinese and Tai" 现代汉语和台语里的助词"了"和"着". *Minzu Yuwen* 民族语文 2,3.

_____. 1979. "On the comparative study of syntax in Sino-Tibetan" 论汉藏语言的比较语法学. *Bulletin of Nankai University* 南开大学学报 1979.4.

_____. 1983. "Some Kam-Sui correspondences of Chinese words in the *yu, xie, zhi, xiao,* and *liu* '*she*'" 汉语遇蟹止效流摄的一些字在侗台语里的对应. *Language Research* 语言研究 1983.1.

_____. 1983. "Chinese-Tai correspondences of the term *bieli*" "别离"一词在汉语台语的对应. *Minzu Yuwen* 民族语文 1983.4.

_____. 1985. *The Sanhong Dialect of Kam* 三红侗语. Nankai University Press 南开大学出版社.

_____. 1986. "The -m/-ŋ alternation in Chinese and Kam-Tai" 汉语和侗傣语里的 -m, -ŋ 交替现象. *Minzu Yuwen* 民族语文 1986.4.

_____. 1989. "A study of 'linking words' in Chinese and Tai" 论汉语台语"关系字"的研究. *Minzu Yuwen* 民族语文 1989.1.

_____. 1989. "The Tai-Ya language of the upper Hong River" 红河上游傣雅语. Language Press 语文出版社.

_____. 1990. "A study of Chinese 'linking words' in the -am, -ap finals of Tai" 台语 -am, -ap 韵里的汉语"关系字"研究. *Minzu Yuwen* 民族语文 1990.2.

_____. 1991. "Some problems related to the genetic relationship between Chinese and Austronesian" 关于汉语南岛语的发生学关系问题. *Minzu Yuwen* 民族语文 1991.3.

_____. 1991. "Chinese-Austronesian initial correspondences" 汉语南岛语声母的对应. *Minzu Yuwen* 民族语文 1991.4.

_____. 1991. "Chinese-Austronesian initial and final consonant correspondences" 汉语南岛语声母及韵尾辅音的对应. *Minzu Yuwen* 民族语文 1991.5.

_____. 1992. "The Tai final -ok is the key to the comparative study of Chinese and Tai" 台语 -ok 韵是汉台语比较的关键. *Minzu Yuwen* 民族语文 1992.6.

_____. 1993. "Substratum correspondences in Chinese-Tai comparative studies" 汉台语比较研究中的深层对应. *Minzu Yuwen* 民族语文 1993.5.

Xu Tongqiang 徐通锵. 1981. "A discussion of some issues in the history of Sino-Xenic language mixture" 历史上汉语和其他语言的融合问题说略. *Yuyanxue Lun Cong* 语言学论丛 7.

_____. 1985. "A study of Matisoff's discussion of historical linguistics and Sino-Tibetan" 马蒂索夫谈历史语言学和汉藏系语言的研究. *Yuyanxue Lun Cong* 语言学论丛 13.

_____. 1991. *Historical Linguistics* 历史语言学. Beijing: Shangwu Yinshuguan 商务印书馆.

Yan Xuequn 严学 . 1979. "A model of change for Chinese Tong lexical inflection" 汉语同族词内部屈折的变换模式. *Zhongguo Yuwen* 中国语文 2.

_____. 1979. "A discussion of cognates and borrowings in Sino-Tibetan" 谈汉藏语系同源词和借词. *Jiang-Han Yuyanxue Cong Kan* 江汉语言学丛刊 1.