THE PHONOLOGY OF KOMPONG THOM CHAM

Robert K. Headley

0. *Introduction*

Cham is a member of the Chamic Group of Austronesian. Other Chamic languages include Chru, Haroi, Jarai, Radê, and Roglai. It is classified into two major dialects: Eastern Cham (spoken by about 30,000 people in the Socialist Republic of Vietnam in the vicinity of Phan Rang) and Western Cham (spoken by about 150,000 people in Vietnam in the vicinity of Châu Đốc and in Tây Ninh Province, and in Cambodia along the Mekong River and its tributaries). This study describes the synchronic and diachronic phonology of the Western Cham dialect spoken in Kompong Thom Province, Cambodia,[1] (KTC). This dialect differs somewhat from the Western Cham dialect of Châu Đốc (CHD) described by Friberg and Hor (1977); it differs substantially from the Eastern Cham of Vietnam (CVN).

The Western Chams apparently represent the descendants of immigrants who left Vietnam after the collapse of the Cham Kingdom in the sixteenth century. The majority of Western Chams are Muslims and use an adoptation of the Arabic script—via Malaysia—to write Cham. Recently, a Latin script has been devised to write Eastern and Western Cham.

1. *The Word*

The native Cham word—which may differ significantly in pattern from borrowed words—has the following syllabic patterns:

MONOSYLLABLES: $(C_3V_2)'C_1(C_2)V_1F$

CV	/la/	['la:]	'leaf'
CVF	/săw/	['sau̯]	'dog'
CVCV	/taho/	['tho:]	'large jar'
	/patɔ/	[pə'tɔ:]	'teach'
CVCVF	/kanɔŋ/	[kə'nɔ:ŋ]	'be angry'

1. This study is based on a corpus of data obtained from a male speaker in his late twenties from Phum Baray in southern Kompong Thom province. The Proto-Chamic reconstructions are mainly from Lee (1965) with some minor modifications, such as *caʔbuay 'mouth' for Lee's *caʔbuai, and a few new reconstructions by the author. Additional Western Cham data are from Friberg *et al.* (1977), and Eastern Cham data are from Blood (1967) and Moussay (1971). Forms between slants are phonemic while forms in italic are generally in the popular orthography of the language.

POLYSYLLABLES: "$C_4V_3C_3V_2F_2$'$C_1V_1F_1$

CVCVCV	/marasa/	["mayə̆'sa:]	'maybe'
CVCVCVF	/panatăŋ/	["panə̆'tàŋ]	'animal'
CVCVFCVF	/patanrăuʔ/	["patə̆n'ɣău̯ʔ]	'make heavy'

In the patterns above: C = any consonant, V = a vocalic nucleus, ' marks primary stress of the following syllable, " marks secondary stress of a following syllable, and F = any final consonant. Borrowed words from Khmer, Vietnamese, Arabic, Malay, Sanskrit, or European sources may show varied syllabic patterns. In some cases, there has been a reshaping of a foreign word into a Cham syllabic pattern. More recent borrowings are less likely to be reshaped than older ones.

/šɔ̆rga/ [šɔ̆rga] 'heaven' (Skt. *svarga*)
/fim/ [fi:m] 'film' (ultimately from French or English 'film')

2. *Register*
Western Cham is a two-register language. The effects of register are seen most obviously on the vocalic nucleus. The High Register (HR) vowels are higher (in terms of tongue height), rather 'breathy' in voice quality, and associated with low pitch. Low Register (LR) vowels are lower in tongue height, often with lower on-glides, rather 'clear' in voice quality, and associated with higher pitch. Friberg and Hor (1977: 18-19), following Gregerson (1976), suggest that the physiological basis for register is the advancement or retraction of the tongue root. They describe First Register (with the tongue root retracted, called Low Register in this paper) vowels as generally lower and tense and Second Register (with tongue root advanced, called High Register in this paper) vowels as generally higher and lax with associated lower pitch and slightly breathy quality.

The High Register developed in syllables which originally, in an earlier stage of Cham,[2] began with voiced consonants. In the dialect of Western Cham described by Friberg and Hor (1977), the High versus Low Register opposition was noted only following stops. Low register vowel nuclei followed originally voiceless stops and high register vowel nuclei followed originally voiced stops. The remaining consonants seemed to be associated with the high register. In the Kompong Thom dialect, original clusters of /h/ + /w, l, y, r, m, n, ŋ/[3] have lost the /h/[4] and have vowel nuclei associated with the low register. In some cases, the original Proto-Chamic (PC) cluster had initial /s/ which must have become /h/ before the Cham Empire broke up, e.g.:

*humã 'field' > /mɨ/ [mᵊɨ] cf. CHD *hamɯ* and CVN *hmu* or *həmu*)
*hanã̍ʔ 'asthma' > /nɨ̃ʔ/ [nᵊɨʔ] (cf. CHD *hanɯc* and CVN *hanɯ'*)

2. Since the traditional Cham script distinguishes between voiced and voiceless stops and it was adopted prior to the eighth century A.D., it is believed that the voiced-voiceless distinction was present in Cham at that time.
3. /h/ + /ñ/ has not been found.
4. In careful speech, the /h/ may be pronounced, but it has been lost in the everyday spoken language.

*sana 'roast' > /na/ [na:] (cf. CHD *hana* and CVN *hna*)
*hureʔ 'vine' > /rɛ̆ʔ/ [ɣɛʔ] (cf. CHD *hrĕk* and CVN *harĕ'*)
*surăʔ 'writing' > /ră̆ʔ/ [ɣaʔ] (cf. CHD *hrak* and CVN *hraʔ* or *sraʔ*)
*sula 'leaf' > /la/ [la:] (cf. CHD *hla* and CVN *hala*)
*haway or huway 'rattan' > /wăy/ [wai̯] (cf. CHD *haway* or CVN *hawĕy*)

The contrast between the reflexes of *w, *l, *y, *r, *m, *n, *ŋ and *h(V) + *w, *l, *y, *r, *m, *n, *ŋ can be seen in pairs like the following.

/ni/ [nᵊiᵀ:] 'bee' < *hunĩ
/nì/ [nìᵃ:] 'this' < *ʔunĩ

/mĭt/ [miᵀt] 'hear' < *hmə̃t
/mìt/ [mìᵃt] 'always' < *(mit)†

/wă̆ʔ/ [u̯aʔ] 'eat' < *huaʔ
/wa̋ʔ/ [u̯à] 'pull down' < *(wăʔ)

/yăw/ [i̯au̯] 'like' < *(hyəu̯)
/yàw/ [i̯àu̯] 'yoke' < *yəu̯

/ră̆ʔ/ [ɣaʔ] 'writing' < *surăʔ
/ra̋ʔ/ [ɣàʔ] 'vein' < *ʔurăt

3. Consonants
The following consonants occur in KTC:

	BILABIAL	DENTAL	PALATAL	VELAR	GLOTTAL
STOPS	p	t	c	k	ʔ
	b	d	j		
NASALS	m	n	ñ	ŋ	
SONANTS	w	l	y	r	
		s			h

/ř, f, š, z/ also occur in borrowed words.

The consonant phonemes of KTC are described in detail below.

/p t c k/ are voiceless, unaspirated bilabial, dental, palatal, and velar stops respectively. They are unreleased in final position. When followed by / ̀/, they are lenis.

/palăy/ 'village' [pə̆'lăi̯]	/paha/ 'thigh' [pha:]
/cìp/ 'Thursday' [čìp]	/tă̆ʔ/ 'chop trees' [taʔ]
/taraʔ/ 'sow rice' [tə̆'ɣ:ʔ]	/tahăw/ 'know' [thau̯]
/matăh/ 'raw' [mə̆'tah]	/ŋăt/ 'be careful' [ŋat]
/cĭm/ 'bird' [čim]	/cakăw/ 'cut with scissors' [čə̆'kau̯]
/tacuiʔ/ 'spit' [tə̆'čui̯ʔ]	/kɛ̆ʔ/ 'to bite' [kɛʔ]
/kapal/ 'thick' [kə̆'pa:l]	/take/ 'horn, antler' [tə̆'ke:]

†Asterisked entries in parentheses ( ) are tentative reconstructions (Ed.).

/palăk/ 'muddy' [pə̆'lã̆k]	/tɔʔ/ 'bottom' [tɔ˞:ʔ]
/tɔ̀ʔ/ 'be at' [tɔ̀ˬ:ʔ]	/ku/ 'Cambodian' [ku˞:]
/kù/ 'kind of taro' [kùˬ:]	/tahu/ 'dry' [thu˞:]
/tahù/ 'bat' [thùˬ:]	/pà/ 'lead' [pà:]
/parà/ 'shoulder' [pə̆'ɣà:]	/tapà/ 'bland' [tə̆'pà:]
/toàh/ 'search for' [toàh]	/tarə̆̀p/ 'of' [tə̆'ɣə̆̀p]
/patə̆̀r/ 'remind' [pə̆'tə̆̀ɯ̆]	/kìŋ/ 'oven' [kì:ŋ]
/karàʔ/ 'rough' [kə̆'ɣà:ʔ]	/carù/ 'medicine' [čə̆'ɣùˬ:]

Modern KTC initial /p t c k/ are reflexes of Proto-Chamic *p, *t, *c, *k and *b, *d, *j, *g. The four voiced PC stops conditioned following high register vowel nuclei.

*patəw 'stone' > /patăw/	*tupaʔ 'straight' > /tapăʔ/
*tubəw 'sugarcane' > /tapə̆̀w/	*durəy 'thorn' > /trùay/
*pagi 'tomorrow' > /pakè/	*bala 'tusk' > /plà/
*ʔadəy 'younger sibling' > /tə̆̀y/	*tiān 'abdomen' > /tean/
*cim 'bird' > /cĭm/	*keʔ 'bite' > /kĕʔ/
*jūʔ 'black' > /cùʔ/	*groh 'to bark' > /krɔ̀h/

The clusters *dl- and *tl become /kl-/.

*tlaw 'three' > /klăw/	*dleh 'tired' > /klə̆̀h/

Voiceless stops are retained in words borrowed from Khmer.

Khm. /phəŋ/ 'inch' > /phə̆ŋ/
Khm. /ckiel/ 'gouge' > /cakial/
Khm. /kmaoc/ 'ghost' > /kamot/
Khm. /tnaot/ 'sugar palm' > /tanot/
Khm. /taə/ 'only' > /təə/.

Some words which contained voiced stops in Middle Khmer were apparently borrowed before these stops were devoiced in Khmer and then subsequently devoiced in KTC. The vowel nuclei in KTC that follow originally voiced Khmer stops are those asociated with the high register.

Khm. /peil/ (*bel*) 'time' > /pèl/
Khm. /koum/ (*gōm*) 'lamp' > /kòm/
Khm. /toap/ (*dāpa*) 'army' > /tɔ̀p/
Khm. /tup/ (*da'pa*) 'stop' > /tɔ̀p/

There are several problems with the reflexes of the final PC stops. Generally, KTC has /-ʔ/ as the reflex for *-p *-t *-c *-k and *-ʔ. Final *-p