

Crawfurd s 1822 Malay of Champa

Graham Thurgood

California State University, Chico

Introduction.¹ In 1822 John Crawfurd, a medical doctor and British civil servant, published as a column in the appendix of a book an 81-item wordlist with the heading Malay of Champa, a designation that is essentially geographical. The existence of the list immediately brings up several basic questions: What language is it? That is, if it is Chamic, precisely which Chamic language is it? Once the language is identified, what can we learn from examining the wordlist? That is, what reliable information does it give us about that language in 1822? In addition, there are a number of minor queries about how to interpret Crawfurd s transcription.

Crawfurd s background. Crawfurd himself was born on the island of Islay west of Scotland in 1783. He trained as a medical doctor at Edinburgh and at twenty he was given an appointment as a medical officer in India s North-West provinces. Crawfurd acquired Malay between 1808 to 1811 in Penang, learning not just the language but the culture. Between 1808 and 1816 he was part of the British presence in Java, including serving as the British Resident at the Court of the Sultan of Jogjakarta. He later served as Resident in Singapore between 1823 and 1826.

Evaluating Crawfurd s historical observations, the eminent Thai historian David Wyatt writes (1967:iv),

Crawfurd was a keen observer who actively sought information, whether from the officials with whom he dealt, foreign residents, Chinese merchants, or simple villagers. The information which he gained is in some places faulty, and on some occasions poorly interpreted, but on the whole it is important, well-organized, and, with a few exceptions, but slightly marred by the author s own prejudices.

Similar praise seems appropriate for Crawfurd s linguistic observations and conclusions.

The Champa wordlist itself occurs in Crawfurd s 1822 *Journal of an embassy to the courts of Siam and Cochin China*, but it is only in his 1852 *Grammar and dictionary of the Malay language; with a preliminary dissertation* that Crawfurd actually discusses it. In the first sentences of his 1852 book, Crawfurd sketches a strikingly modern account of the geographical distribution of the Malayan languages (1852:a):

A certain connexion, of more or less extent, is well ascertained to exist between most of the languages which prevail from Madagascar to Easter Island in the Pacific, and from Formosa, on the coast of China, to New Zealand. It exists, then, over two hundred degrees of longitude and seventy of latitude, or over a fifth part of the surface of this earth.

He then continues describing the geographical spread of Malayan

as consisting of the innumerable islands of the Indian Archipelago, from Sumatra to New Guinea of the great group of the Philippines of the islands of the North and South Pacific and of Madagascar.

There is no doubt that Crawford had recognized the Formosan component of Austronesian or Malayan, as he terms it (1852:cxxxiii). Citing Klaproth's *Asia Polyglotta* as the source of his forms, Crawford notes that the following are of undoubted Malayan origin:

Formosa	English	Formosa	English
wato	stone	sat, s at	one
mata	eye	rauha	two
lima	hand	tauro	three
tangira	ear	hipat	four
wa a	fruit	rima	five
ap i	'fire	n um	six
aulong	man	pitu	seven
alak	son	audim	black
reia	joy, pleasure	paule	white

However, it is not Crawford's Formosan identifications but his Champa wordlist that is of interest here.

The list itself. Some preliminary comments on the list itself are in order. The 1822 volume was first presented by Crawford to the Indian Government on his return to Calcutta, and later published with only formatting changes in 1828 as a quarto volume, and then again in 1830 as two quarto volumes. The 1967 version is identical to the 1822 version, but accompanied by an excellent introduction by David Wyatt, the source of this information.

Although Crawford does not say so, the list is the same list he published earlier in 1822 as part of the appendix at the end of Crawford's *Journal of an embassy to the courts of Siam and Cochin China*.² Aside from the quibble that the 1822 list has 79 rather than 81 items, the 1822 list, rather unambiguously labeled the Malay of Champa, is the one referred to in Crawford 1852. Without exception, all the items appearing in the 1852 discussion also appear on the 1822 list, including the inexplicable but instructive spelling of thousand as *rilau* (rather than the expected *ribau*). And, with the exception of

four items, they occur in exactly the same form. His earlier *plu ten* has inexplicably been reduplicated, producing the euphonious but strange *plu plu*. Incidentally, the word *gunong* was incorrectly placed in the column for island; this is undoubtedly a printer's error, as Crawfurd would have known the Malay word.

The vowel representations (and their apparent emendations in 1852) are intriguing. The vowel in his earlier *mus gold* has been replaced, producing the more anticipated *mas*. Whether this is simply Crawfurd fixing a printer's error in the earlier list, or whether the *-u-* represented a high, mid, central vowel, that Crawfurd inappropriately emended *-u-* to *-a-* we will never know. The earlier form, however, is the more interesting as modern Cham has a high, mid, central vowel in the word. Crawfurd's earlier *preak silver* has been replaced by *pr ak silver*, a change that is fascinating, as the word mostly likely contained a high, mid central vowel. Finally, the final vowel in his earlier *naharai sun* with the expected vowel reflex has been replaced by *nahari sun* with an unexpected vowel reflex. Again, whether Crawfurd is again fixing an earlier printer's error, or whether he has erroneously emended *-ai-* to *-a-*, again we will never know, although I suspect the latter.

As for Crawfurd's transcription, it matches up well with modern Cham and seems reliable, except for his transcription of final *-h* and final *-ʔ*. which appears ambiguous between a final glottal stop and a final *-h*. In two of the six words in which final *-h* are expected, Crawfurd has nothing; in the remaining two, the final *-h* appears.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawfurd's Champa Malay	
ratus	*ratus-f	ratuh	ratuh	ratũh	ratu	hundred
beras	*bra:s	bra	prah	prah-l	bra	rice(husked)
---	*masuh	masuh	masruh-r	mithuh	musu	fight (war)
se-puluh	*pluh	sa pluh	hapluh	plũh	plu ten	
habis	*ʔabih	abih	pih	apih	amubeh	all; finished, done
panah	*panah	panãh	panih	panĩh	panah	shoot (bow); a bow hunting bow
---	*ʔakoʔ	akoʔ	kšʔ	akšʔ	akoh	head
(ə)ma(?)	*mɛʔ-vf	---	mɛʔ	mɛʔ	meh	mother
langit	*lanjit	lanji:ʔ	lanjiʔ	lanjiʔ	langi	sky

As for the glottal stop, in two of the three words in which a glottal stop is expected, it appears marked with *-h*; in the other word, it is missing. It certainly was not unknown to use a final *-h* for a glottal stop. Some fifty years later, McNair (1878:7), for example, wrote concerning the Malay of Perak Perak pronounced as though spelt Payrah for a word undoubtedly

pronounced with a final glottal stop. The form -u- in the word *amubeh* only makes sense as some sort of printer's error for, most likely, *ambeh*. However, this means that the symbol -h is ambiguous between final -h and glottal stop.

What language is it? The first major question is precisely what language is contained in Crawfurd's wordlist. Certainly, Crawfurd himself considered it the Malay of Champa, and writing later in 1852 (cxxxix), he stated:

The only part of the continent of Asia, the Malay peninsula excepted, in which the Malays have settled, and to which their language has extended, is Kambodia, correctly Kamboja, which appears to be a Malay word. In that country they have established a little independent principality called Champa, well known both in Malay and Javanese story. It was from a merchant of this country trading with Singapore, that I received a short list of 81 words of the language of Champa.

Even had Crawfurd not headed the 1822 list MALAY OF CHAMPA, the source would have been readily recognizable. Its close relationship to the Malay languages is evident from the innovative numerals and, possibly, its sharing the innovated h- in the word for day with Malay.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawfurd's Champa Malay	
tujuh	*tujuh	tijuh	taçuh	taçũh	tuju	seven
delapan	*dua-lapan	---	tapån	ṭalipån	dalapan	eight
sembilan	*samilan	---	samlån	samilan	samilan	nine

The semantics of the forms, discussed by Blust (1981:467, fn. 5), derives from the use of the pointing finger during counting for seven and from roughly two taken from ten (< *dua-alap-an) and one taken from ten (< *se-ambil-an).

The evidence that it is Chamic comes from a configuration of features present and features absent. Like Chamic, it does not share the Malay innovation of *satu* 'one' and *tiga* 'three', nor does it share the change of *-uy > -i, as in *apuy > Malay *api* and *babuy > Malay *babi*: found in Malay dialects. Thus, despite being close to Malay, Chamic does not fit within Adelaar's proto-Malayic (1988), probably better termed proto-Malayan, which has *api and, presumably, *babi.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawfurd's Champa Malay	
(satu)	*sa	sa	sa;ha	tha	satu	one
(tiga)	*klɔw	tləu	klau	klăw	klao	three
api	*?apuy	apui	pui	apuy	apoi	fire
babi	*babuy	babui	papui	papuy	baboi hog	wild pig

The wordlist shares certain features with Chamic: First, the diphthongization of the PAN high vowels *-i and *-u in word-final position to PC *-ɛy and *-ɔw, which then became Phan Rang Cham *-ăy and -ăw.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawford s Champa Malay	
matahari	*ia hurey	ia hurəi	ea hray	ýa harăy	nahari; naharai	sun (< 'day')
besi	*bisey	pisəi	pasay	pithăy	basai	iron
lakilaki	*lakey	lakəi	lakay	likăy	lakai	male; person
negeri	*n/lagar	laga	naċar -i	---	nangrai	country; city; area
batu	*batɔw	patəu	patau	patăw	batao	stone
tebu	*tabɔw -v	tubəu	tapau	tapăw	tabao	sugarcane

Second, it shares the loss of all homorganic nasal plus stop combinations except in borrowed words.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawford s Champa Malay	
empat	*pa:t	pa:ʔ	paʔ	paʔ	pak	four
sembilan	*samilan	---	samlăn	samilan	samilan	nine

And, third, it shares the presence of several words, found in Chamic but not commonly found elsewhere. Some were incorporated into pre-Chamic from Mon-Khmer sources before the formation of Proto-Chamic; some are Sanskrit borrowings, e.g., ten thousand and plow; and some may turn out to be from other sources.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawford s Champa Malay	
---	*lamɑ:n	lumăn	lamin	limin	lamun	elephant
---	*ge	ge	ķe	ķe	ge	boat
kain	*khan	khat	khăn	khăn	kan	cloth;blanket
---	*cuɓuai;	cu ɓuai;	caɓuai	caboy	chabui	lips;gums
---	*tuɓuai	cubuai			mouth	
---	*sa xsit	---	hasit -f; sit -f	asit -i; sit; sīt	asit	little
(ə)ma(?)	*mɛʔ -vf	---	mɛʔ	mɛʔ	meh	mother

Having established that the language is Chamic, it is also apparent that it is Cham, because of the excellent match of lexical items. The identification of the list as specifically Phan Rang Cham is not from its linguistic features *the*

linguistic features do not distinguish it from Western Cham, but from our knowledge that by 1822, the Western Cham speakers had split off from the Phan Rang Chams, moving westward from eastern Vietnam after the Cham federation collapsed in the fifteenth century (Headley, 1991).

What can we learn from it? Having identified the language as Phan Rang Cham, what can we learn about Phan Rang Cham and, possibly Chamic, from this wordlist?

Fossilized morphology. The word *naharai / nahari* sun seems to contain fossilized morphology. The initial nasal appears to be a remnant of the PAN *ni genitive, as found in the Fijian *mata ni siga* sun i.e., eye of the day and in *mata ni ari* sun i.e., eye of the day of the more closely related Toba Batak.

		N.	W.	PR	Crawfurd s
Malay	PC	Roglai	Cham	Cham	Champa Malay
bulan	*ia bula:n	ia bilat	eaplan	pilan	bulan moon;month
matahari	*ia hurey	ia hurai	eahray	ya haray	nahari; sun < day naharai

The presence of this form for sun leads to the reexamination of the Chamic words for both sun and moon. Although dictionaries often implicitly treat the first element of both as if it meant the homophonous water, it is more likely that Aymonier and Cabaton (1906) were right relating the first element to *yaŋ* deity; it is also quite possible that this element is a remnant of fossilized morphology. Quite interesting for this reason are the forms in Rade, *yaŋ hrue* sun, and a variant of this which occurs in Aymonier and Cabaton's Cham dictionary.

Malay influence. Of general interest is the large percentage of apparently Malay words in Crawfurd's list, including a number of known borrowings. That is, many of the words in the list occur in Malay but not in other Chamic sources. The temptation to assume that these are simply unattested Chamic forms needs to be resisted. In some cases, this is clear from the structure: *manis* is suspicious because of its final -s (if -s were retained in this word, why not in *ambeh* finished?), *bintang* and *anjing* are suspicious because the nasal component of homorganic stop plus nasal combinations is otherwise generally lost throughout Chamic.

The last three forms have been included in this list, because despite the existence of related Chamic forms, the Crawfurd forms are more closely related to the Malay than the Chamic. The forms for sweet represent a doublet, with Malay having one of the reflexes and most of Chamic the other, although Written Chamic contains both members of the doublet.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawford's Champa Malay	
pedang	---	---	---	---	padang	sword
bapak	---	---	---	---	pak	father
laut	---	---	---	---	laut	sea, ocean
si-apa	---	---	---	---	siapa	who?
sungai	---	---	---	---	sungai	river
lada	---	---	---	---	lada	pepper
tenun	---	---	---	---	tanun	to weave
banyak	---	---	---	---	banyak	many
gunung	---	---	---	---	gunong	mountain
elok	---	---	---	---	elok	good
handsome; beautiful						
sutera	---	---	---	---	stro	silk
bintang	---	---	---	---	bintang	star
tembaga	---	---	---	---	tambaga	copper; brass
añjij	---	---	---	---	anjing	dog
manis	*mamih	mumih	---	mimih	manis	sweet
rimau	*rimo:ŋ	lumōŋ -i	ramoŋ	rimoŋ	rimao -f	tiger
timah	*tamra?	tumra:?	tamra?	tāmra?	tima	tin; lead

The existence of such a high percentage of Malay suggests the possibility of a heavy Malay influence in the Phan Rang Cham of 1822, or at the very least in the speech of the Champa merchant who supplied this list. However, various entries scattered throughout Aymonier and Cabaton's 1906 Cham dictionary show a similar Malay influence but without the precise date supplied by Crawford. As a corollary, the presence of an item on this list cannot be taken as conclusive evidence that it was ever Chamic.

Naturally, the vast majority of the forms are well-attested Chamic words, some found in Malay and some not.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawford's Champa Malay	
---	*pɔ	po	pɔ title	po	po	master; lord
---	*prɔŋ	prok	pruŋ -v	prɔŋ	prong great	big
---	*kuməy	kuməi- ⁿ	kamay	kaməy	komai	female, woman
---	*ʔako?	ako?	kɔ?	akɔ?	akoh	head
dua	*dua	dua	ʔoa	ʔwa	dua	two
empat	*pa:t	pa:?	pa?	pa?	pak	four
lima	*lima	lumã	lami	limi	limo	five
enam	*nam	nãm	nãm	nãm	num	six

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawfurd s Champa	(continued) Malay
sebelas	*sa pluh sa	sa pluh sa	ha pluh sa	tha plüh tha	plu-sa	eleven
dua- puluh	*dua- pluh	dua- pluh	doa- pluh	ṭwa- plüh	dua- plu	twenty
ribu	*ribəw	rubəu	rapau	ripǎw	rilao	thousand
jahat	*jahat	---	---	çha?	jahat	bad; wicked
pahit	*phit	phi:?	phi?	phi?	pahit	bitter; bile
mata	*mata	mata	mata	mita	mata	eye
sedikit	*dikit	tiki:?	taki?	taki?	sadikit	few; little
emas	*ama(:)s	māh	mih	mǎh	mus, mas	gold
ikan	*ika:m	ika:t	kan	ikan	ikan	fish
sini	*tinī	tinī	ni	ni	nao -v here	here
ulun	*hulun	hulut	hulɪn I	halün	alun I	slave; I(polite)
orang	*ura:ŋ	urak	raŋ	uraŋ	orang	person; someone
ayer	*iar -f	ia	ea -f	ier	aya	water (fresh)
atas	*ata:s	ata	tah	atǎh	adai	far; above; long
buat do	*buat	buã?	---	---	buat do	do; work
kerbau	*kabaw	kabau	kabau	kapaw	kubao	water buffalo
perak	*pirak-lf	paria? (m)	parea? (m)	paryǎ? (m); pirak -f	preak; priak silver	silver; money; white

The chronology of sound changes. Finally, and potentially the most interesting, the list may provide some suggestive evidence about the chronology of certain sound changes, particularly the devoicing of the voiced obstruents of Phan Rang Cham. Certainly the devoicing had occurred by 1901 as Cabaton (1901:68) mentions that *La sonore malaise se change en sourde*. But, in Crawfurd s 1822 list the voiced obstruents are recorded as voiced! It looks as though the change must have taken place between 1822 and 1901, but as nice as it would be to know with such precision when the change took place, the Crawfurd data cannot be trusted in this instance. Crawfurd obtained the list himself, and it is quite conceivable that as a native speaker of English he would have heard and recorded voiceless unaspirated stops, particularly if accompanied by breathy phonation, as voiced stops. Thus, the list brings no real evidence to bear on the dating of obstruent devoicing.

Questions remaining. Of the eight items remaining on Crawfurd s list, seven could not be related to items in Chamic, nor in Austronesian.

Malay	PC	N. Roglai	W. Cham	PR Cham	Crawford s Champa Malay	
---	---	---	---	---	kaoya	you; thou
---	---	---	---	---	mandao be	copula
---	---	---	---	---	naoya	was
---	---	---	---	---	boat	will
---	---	---	---	---	pala	near
---	---	---	---	---	maya	far, distant
---	---	---	---	---	taggo	below
---	---	---	---	---	naweh	give

The last item *naweh* give might very well be related to the Javanese *ngu-w h* give, a Javanese Ngoko form, that is, used in informal speech. Whether the other items ultimately turn out to be interesting remains to be seen.

Conclusions. On the basis of linguistic criteria, Crawford s phonetically quite reliable Malay of Champa is identifiable as Cham, and on the basis of other non-linguistic evidence is further identifiable as Phan Rang Cham. The word for sun is particularly interesting for the apparent fossilized remnant of the PAn *ni genitive, which it contains. More generally, the examination of the list shows a heavy Malay component in 1822 Phan Rang Cham, at least among certain classes, making it impossible to assume that a word on Crawford s list is Cham without collaborating evidence from elsewhere in Chamic. Finally, Crawford s list fails to show any clear indication of obstruent devoicing, leaving us without any evidence for the dating of that change.

NOTES

¹I wish to thank several people for their feedback on earlier versions of this paper: Bob Blust, George Grace, Elzbieta Thurgood, and Isidore Dyen. I shall be astonished if all my errors should prove minor and grateful to readers for their corrections. This work is based upon work supported by the National Science Foundation under Grant No. SBR-951211011.

The following conventions are used in this paper. The reconstructions of proto-Chamic (PC) are my own (Thurgood 1999). PR is Phan Rang Cham. A subscribed dot under a voiceless obstruent marks a second register syllable; that is, among other things, the formerly voiced consonant is now voiceless and that the following vowel has a low tone in Phan Rang Cham and is in the breathy voiced second register in Western Cham. In this and subsequent tables, borrowings are marked with superscripted ^x and borrowings that predate the formation of PC with ^{x*}. Irregularities are marked with a minus; thus, -v indicates an unexpected vowel, -f an unexpected final, -l length, -ⁿ nasalization, and so on.

²This volume was first presented by Crawford to the Indian Government in 1822 on his return to Calcutta, and later published with only formatting changes in 1828 as a quarto volume, and then again in 1830 as two quarto volumes. The 1967 version is identical to the 1822 version, but accompanied by an excellent introduction by David Wyatt, the source of this information.

REFERENCES

- Horne, Elinor C. 1961. *Beginning Javanese*. New Haven and London: Yale University Press.
- Adelaar, K.A. 1985. *Proto-Malayic, the reconstruction of its phonology and parts of its lexicon and morphology*. Alblasterdam: Offsetdruk-kerij Kanters B.V.
- Aymonier, tienne Fran ois and Antoine Cabaton. 1906. *Dictionnaire cham-fran ais*. PEFE0 7. Paris, Leroux.
- Blust, Robert A. 1981. The reconstruction of Proto-Malayo-Javanic: an appreciation. *Bijdragen tot de taal-, land-, en volkenkunde* 137.4:456-469.
- Cabaton, Antoine. 1901. *Nouvelles recherches sur les Chams*. Publications de L' cole Fran ais D'Extr me-Orient. Volume II.
- Crawfurd, John. 1822 [1967]. *Journal of an embassy to the courts of Siam and Cochin China* [reprinted]; with an introduction by David K. Wyatt. Kuala Lumpur, London, New York: Oxford University Press.
- Crawfurd, John. 1852. *Grammar and dictionary of the Malay language; with a preliminary dissertation*. Volume I: Dissertation and Grammar. London: Smith, Elder, and Co., 65, Cornhill.
- Headley, Robert K. 1991. The phonology of Kompong Thom Cham. *Austroasiatic Languages: Essays in honour of H. L. Shorto*. Edited by Jeremy H. C. S. Davidson. University of London, School of Oriental and African Studies: 105-121.
- McNair, J. F. 1972 [1878]. *Perak and the Malays*. Kuala Lumpur: Oxford University Press.
- Thurgood, Graham. 1999. *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change*. Oceanic Linguistics Special Publication No. 28. Honolulu: University of Hawai i Press.
- Wyatt, David K. 1967. Introduction. *Journal of an embassy to the courts of Siam and Cochin China* [reprinted], by John Crawfurd. Kuala Lumpur, London, New York: Oxford University Press. 4pp.

gthurgood@csuchico.edu

The dedication of the book immediately catches the eye.

To
THE BARON ALEXANDER VON
HUMBOLDT.

Sir,

I dedicate this Work to you, on account of the high respect which, in common with the rest of the world, I entertain for yourself; and in testimony of my veneration for your distinguished brother, whose correspondence on the subject of my labours I hold in grateful recollection.

I am, with great esteem,

Your faithful Servant,

J. CRAWFURD

¹I wish to thank several people for their feedback on earlier versions of this paper: Bob Blust, George Grace, Elzbieta Thurgood, and Isidore Dyen. I shall be astonished if all my errors should prove minor and grateful to readers for their corrections. This work is based upon work supported by the National Science Foundation under Grant No. SBR-951211011.

The following conventions are used in this paper. The reconstructions of proto-Chamic (PC) are my own (Thurgood 1999). PR is Phan Rang Cham. A subscribed dot under a voiceless obstruent marks a second register syllable; that is, among other things, the formerly voiced consonant is now voiceless and that the following vowel has a low tone in Phan Rang Cham and is in the breathy voiced second register in Western Cham. In this and subsequent tables, borrowings are marked with superscripted ^x and borrowings that predate the formation of PC with ^{x*}. Irregularities are marked with a minus; thus, -v indicates an unexpected vowel, -f an unexpected final, -l length, -ⁿ nasalization, and so on.

²This volume was first presented by Crawford to the Indian Government in 1822 on his return to Calcutta, and later published with only formatting changes in 1828 as a quarto volume, and then again in 1830 as two quarto volumes. The 1967 version is identical to the 1822 version, but accompanied by an excellent introduction by David Wyatt, the source of this information.