

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization  
Organisation internationale de normalisation  
Международная организация по стандартизации

**Doc Type:** Working Group Document  
**Title:** Proposal to Encode Western Cham 2021  
**Source:** Alberto Pérez Pereiro, Leb Ke, Jorge López Cortina & Andrew Cunningham  
**Action:** For consideration by JTC1/SC2/WG2  
**Date:** 2022-02-12

## 0.0 Introduction

The Western Cham language spoken in Cambodia and in the Mekong Delta in Vietnam is closely related to but not completely mutually intelligible with both the Eastern Cham language spoken on the central coast of Vietnam and the Jarai language spoken in the highlands of Vietnam and Cambodia. These Chamic languages have a common origin in the languages spoken on the coast of what is today Vietnam in the 1<sup>st</sup> century AD, and have taken their present forms as a result of migrations from this original homeland to the Annamite mountains, the territories that are now the Kingdom of Cambodia, and the Delta of the Mekong river in southern Vietnam (Brunelle and Thurgood 2015).

While the Cham people acquired writing in Sanskrit, like many other southeast Asian peoples, they did eventually begin to write in Cham itself, with the earliest inscription known, the Đông Yên Châu inscription dated to the 4<sup>th</sup> century AD representing the earliest extant example of writing in any Austronesian language (Coedès 1939). This script survives today in two forms, one used in Vietnam for Eastern Cham and another used in Cambodia for Western Cham. While the two scripts are clearly related, there are differences in their usage, and as they are used to write languages that have diverged considerably from each other, it is advisable to have a separate Unicode Block for Western Cham. This proposal seeks to be as inclusive as possible of all of the contemporary and historic varieties of use of the Western Cham script.

## 1.0 Western Cham Scripts and Orthography

There are two scripts currently in common use to write the Western Cham language. One of these, 'Jawi' is a modification of the Malay Jawi script which is itself a modification of the Arabic script. The other script, with which this proposal is concerned, is typically referred to as Ka-Kha after the first two letters of the script. It is an Indic writing system, clearly related to other Southeast Asian writing systems such as Khmer, or Thai.

There is no official regulator of the Cham language, neither of its grammar or its orthography. While the Cambodian government does regulate the use of the Khmer language and the languages of the nation's indigenous minorities, those minorities with origins outside of Cambodia, namely the Chinese, Vietnamese and Cham, are permitted to write their languages as they choose. Cham

writing in the Ka-Kha script exhibits slight variations of orthography and spelling conventions, although for the most part any literate Cham person can decipher texts written in an orthography other than their own.

While until recently, the use of the Ka-Kha script was restricted to the Imam San religious group, there has been a revival of interest in this native script in recent years in other communities. Moreover, research on Cham graffiti at Angkor Wat has proven conclusively that the script was in use outside of the Imam San community up until the Cambodian Civil War and the Pol Pot régime in the 1970's (Leb Ke 2018). (See [Figure 1](#))

This proposal is for an encoding that permits for all of the orthographies currently in use. While some writers do not use some of the letters or characters in this proposal, any characters they do use are found in this proposal. The orthography used varies from community to community with different villages and lineages of teachers having arrived at slightly different usages of the script to represent their language. For this reason, it is important for the encoding of the Western Cham language to be as inclusive as possible in order to accommodate the many variations in which historical and contemporary texts are written.

This proposal corresponds to the recommendations of the Cham Language Advisory Committee (CLAC), an advisory body established in 2011 with the oversight of the Ministry of Education, Youth and Sport of the Cambodian government. The members of the CLAC are men and women from different regions of the country where Cham is spoken and from different religious communities. This is important as some variation in orthography is due to dialect, and because, while all Cham people in Cambodia are Muslims, the particular form of Islam that they practice influences how they write and even speak their language, with the Imam San community using the Ka-Kha script and other Muslim groups habitually using the Arabic-based Jawi script, although interest in using the Ka-Kha script has steadily grown.

The CLAC recommended orthography is the most expansive used in order to write the language. This is to say that all other spelling conventions found in the country are subsets of the CLAC recommended system, and writers who prefer different orthographies would be able to use the encoding in this proposal to write in the way they see most fit. Moreover, the CLAC orthography has been the most widely taught in the country and is the only orthography that is taught outside of the Imam San religious community.

### **0.1 Encoding Principles**

As with other Indic scripts, each Cham character is pronounced with an inherent vowel.

1. For most characters, this is [aʔ].
2. For aspirated, lax consonants this is [a:] with a lower tone and a breathy phonation.
3. For each nasal consonant there are two possible inherent vowels and this is indicated by using variant characters. They are [iʔ] for the basic form of the character and [aʔ] for the modified form. Both of these forms are displayed in the Nasal Column of Section 1.0.

Also similar to other Brahmi-based scripts, Cham vowel symbols are attached to their consonants. These symbols may appear above, below, after, or before the consonant to which they are attached, and multiple vowel symbols and diacritics may be attached to a consonant at the same time. In addition to this, there are two subscript consonants, R and L, which are attached to the main consonant much as the vowel signs are.

### **0.2 Punctuation**

Cham uses a number of punctuation marks including the Danda, the Double Danda and the Triple Danda. The Double Danda corresponds to a full stop, and the Triple Danda is used to mark the end of chapters or sections. The Danda can be used to separate words or phrases, although this is usually accomplished by inserting a space between the elements.

A reduplication symbol based on the Arabic numeral 2 is also used. Religious texts in particular make use of other symbols in order to mark different sections of text. Periods may be used in abbreviations and mathematical expressions. Commas, hyphens, question marks, exclamation marks and other symbols may be used in a manner similar to European languages according to the preference of the author.

### **0.3 Spacing and Line Breaks**




















Individual words are not separated by spacing in Cham. Spacing serves to separate phrases and to offset certain words. When typing, individual words should be separated by a zero width space (U+200B). This usage is effectively the same as in Khmer. When necessary, words can be broken at a syllable boundary.

## 1.0 Cham Consonants

The following charts shows the Cham consonants as well as their final variants placed below. There exist other variants that are found in manuscripts and preferred by certain writers not found in this table which will be explained and included in this proposal.

	Occlusive							Sonorant	Fricative		Special
	Tense		Lax							Tense	Lax
	Unaspirated	Aspirated	Unaspirated	Aspirated	Nasal		Voiced				
Glottal	ꨀ								ꨁ		
IPA	ʔa?								ha?		
Velar	ꨂ	ꨃ	ꨄ	ꨅ	ꨆ	ꨇ		ꨈ			
	ꨉ		ꨊ		ꨋ			ꨌ			
IPA	k̚a?	k̚ʰa?	ka?	k̚ʰa:	ŋi?	ŋa?		ɣa?			
Palatal	ꨍ	ꨎ	ꨏ	ꨐ	ꨑ	ꨒ	ꨓ	ꨔ			
	ꨕ							ꨖ			
IPA	ç̚a?	ç̚ʰa?	ca?	c̚ʰa:	ɲi?	ɲa?	ʃa?	ja?			
Dental	ꨗ	ꨘ	ꨙ	ꨚ	ꨛ	ꨜ	ꨝ	ꨞ	ꨟ	ꨠ	ꨡ
	ꨛ				ꨛ			ꨞ	ꨟ	ꨠ	ꨡ
IPA	ʈa?	ʈʰa?	ta?	t̚ʰa:	ni?	na?	ɗa?	la?	ʂa?	za?	nin
Labial	ꨣ	ꨤ	ꨥ	ꨦ	ꨧ	ꨨ	ꨩ	ꨪ			ꨫ
	ꨬ	ꨭ	ꨮ		ꨯ			ꨰ			
IPA	p̚a?	p̚ʰa?	pa?	p̚ʰa:	mi?	ma?	ba?	va?			p̚a?

## 1.1 Glyph Placement

Syllable	Glyphs	Code Point Sequence
ka:		1E200 + 1E22E
kiʔ		1E200 + 1E22F
ki:		1E200 + 1E230
kaj		1E200 + 1E231
kiʔ̃		1E200 + 1E232
ki:		1E200 + 1E233
ke:		1E200 + 1E234
ka:j		1E200 + 1E235
kao		1E200 + 1E236
ka:o		1E200 + 1E234 + 1E236
ki:		1E200 + 1E22E + 1E233
keə		1E200 + 1E238
koə		1E200 + 1E23B
klaʔ		1E200 + 1E23A
kʸaʔ		1E200 + 1E239
kaŋ		1E200 + 1E241
kam		1E200 + 1E249
kiŋ		1E200 + 1E233 + 1E241
kim		1E200 + 1E233 + 1E249



























## 1.2 Representation of Tenseness in Cham

A characteristic of the Cham language is the distinction of tense and lax consonants. This is a complex phenomenon that has consequences for the phonation and tone of the following vowels (Brunelle 2005; Ueki 2011). If one compares the table of Cham consonants to the Devanagari, it is apparent that the voiceless consonants are used to represent tense sounds and the voiced consonants are implemented for lax sounds. This correspondence mirrors that found in the Khmer

distinction between first and second series consonants, or the Tibetan distinction of high and low tone consonants.

For the majority of sounds, there are different characters representing tense and lax versions. For those characters without the tense/lax counterpart, a diacritic mark, SIGHN SALIH SAP, is used to change this quality. The use of diacritics to change tenseness sometimes results in obligatory ligatures.

In IPA transcription, tense consonants will be indicated by double bars below ([kaʔ] vs. [k͡aʔ]).

Character	Word	IPA	+ SALIH SAP	IPA
NGUE		[ŋiʔ]		[ŋ͡iʔ]
NGA		[ŋaʔ]		[ŋ͡aʔ]
NHUE		[ɲiʔ]		[ɲ͡iʔ]
NHA		[ɲaʔ]		[ɲ͡aʔ]
NUE		[ɲiʔ]		[ɲ͡iʔ]
NA		[naʔ]		[n͡aʔ]
MUE		[miʔ]		[m͡iʔ]
MA		[maʔ]		[m͡aʔ]
YA		[jaʔ]		[j͡aʔ]
RA		[ɣaʔ]		[ɣ͡aʔ]
LA		[laʔ]		[l͡aʔ]
VA		[uaʔ]		[u͡aʔ]
HA		[haʔ]		[h͡aʔ]




















## 2.0 Characters to Be Included in the Unicode Block

Code	Character	Name	IPA	Notes
<b>Consonants</b>				
1E200	ꨀ	KA	kaʔ	
1E201	ꨁ	KHA	kʰaʔ	
1E202	ꨂ	GA	gaʔ	
1E203	ꨃ	GHA	kʰa:	
1E204	ꨄ	NGUE	ŋiʔ	
1E205	ꨅ	NGA	ŋaʔ	
1E206	ꨆ	CHA	ɕaʔ	
1E207	ꨇ	CHHA	ɕʰaʔ	
1E208	ꨈ	JA	caʔ	
1E209	ꨉ	JHA	cʰa:	
1E20A	ꨊ	NHUE	ɲiʔ	
1E20B	ꨋ	NHA	ɲaʔ	
1E20C	ꨌ	NHJA	ʃaʔ	
1E20D	ꨍ	TAK	taʔ	
1E20E	ꨎ	THA	tʰaʔ	
1E20F	ꨏ	DA	daʔ	
1E210	ꨐ	DHA	tʰa:	
1E211	ꨑ	NUE	niʔ	
1E212	ꨒ	NA	naʔ	
1E213	ꨓꨔ	SIGN NOEN	nɪn	Obligatory ligature used to write the word [nɪn] meaning “that”.
1E214	ꨕ	DDA	ɗaʔ	
1E215	ꨖ	PA	paʔ	

# Proposal to Encode Western Cham 2022










































































































1E216	𑜋𑜧	PA TNAW	paʔ	Used obligatorily in the words [paʔ] meaning “four”, and [po] meaning “God”.
1E217	𑜋𑜧𑜨	PA PRONG	paʔ	Used in the word Jibril (The name of the Archangel Gabriel in Arabic)
1E218	𑜋𑜧𑜨𑜨	PHA	p <sup>h</sup> aʔ	
1E219	𑜋𑜧	BA	paʔ	
1E21A	𑜋𑜧𑜨	BHA	p <sup>h</sup> a:	
1E21B	𑜋𑜧𑜨𑜨	MUE	miʔ	
1E21C	𑜋𑜧𑜨𑜨	MA	maʔ	
1E21D	𑜋𑜧𑜨𑜨	BBA	baʔ	
1E21E	𑜋𑜧𑜨𑜨𑜨	YA	jaʔ	
1E21F	𑜋𑜧𑜨𑜨𑜨	RA	yaʔ	
1E220	𑜋𑜧𑜨𑜨𑜨	LA	laʔ	
1E221	𑜋𑜧𑜨𑜨𑜨	VA	uaʔ	
1E222	𑜋𑜧𑜨𑜨𑜨	SSA	zaʔ	
1E223	𑜋𑜧𑜨𑜨𑜨	SA	saʔ	
1E224	𑜋𑜧𑜨𑜨𑜨	HA	haʔ	
1E225	𑜋𑜧𑜨𑜨𑜨	A	ʔaʔ	
1E226	𑜋𑜧𑜨𑜨𑜨𑜨	A DHA	ʔaʔ	Variants of A used by some writers for proper names and words derived from Arabic. These characters may co-occur in the same text.
1E227	𑜋𑜧𑜨𑜨𑜨𑜨	A BA	ʔaʔ	
1E228	𑜋𑜧𑜨𑜨𑜨𑜨	A U	ʔaʔ	
1E229	𑜋𑜧𑜨𑜨𑜨𑜨	I	ʔiʔ	
1E22A	𑜋𑜧𑜨𑜨𑜨𑜨	U	ʔuʔ	Also [ʔoʔ]
1E22B	𑜋𑜧𑜨𑜨𑜨𑜨	E	ʔeʔ	
1E22C	𑜋𑜧𑜨𑜨𑜨𑜨	AI	ʔa:j	
1E22D	𑜋𑜧𑜨𑜨𑜨𑜨	OH	ʔoh	
<b>Vowel Signs</b>				
1E22E	𑜋𑜧𑜨𑜨𑜨𑜨	SIGN AA	a:	



1E22F		SIGN I	i	
1E230		SIGN II	i:	
1E231		SIGN AI	aj	
1E232		SIGN OEK	ɨʔ	
1E233		SIGN OE	ə	
1E234		SIGN E	e	
1E235		SIGN AAY	a:j	
1E236		SIGN AW	ɑ	
1E237		SIGN U	u	
<b>Subscript Consonants</b>				
1E238		SIGN Y	ɛə	Traditionally considered a subscript consonant. It is used for the diphthong [ɛə].
1E239		SIGN R	ɣ	A conjunct consonant used to write consonant clusters such as [kɣ], [tɣ], or [pɣ].
1E23A		SIGN L	l	A conjunct consonant used to write consonant clusters such as [kl], [tl], or [pl].
1E23B		SIGN V	oə	Traditionally considered a subscript consonant. It is used for the diphthong [oə].
1E23C		SIGN SALIH SAP	tense «» lax	Used to change tense consonants into lax ones and vice-versa. The use of this character result in several ligatures. Not all writers distinguish between this character and SIGN L.
<b>Final Consonants</b>				
1E23D		FINAL K	_ʔ	
1E23E		SIGN FINAL K	_ʔ	Optional ligature equivalent to the FINAL K.
1E23F		FINAL G	_kʔ	
1E240		FINAL NG	_ŋ	
1E241		SIGN FINAL NG	_ŋ	Indicates a preceding short vowel.

1E242	ꨀ	FINAL CH	<u>ik</u>	
1E243	ꨁ	FINAL T	<u>t</u>	
1E244	ꨂ	FINAL N	<u>n</u>	
1E245	ꨃ	FINAL P	<u>p</u> / auʔ	
1E246	ꨄ	FINAL PH	<u>p</u> / f	
1E247	ꨅ	FINAL B	<u>p</u>	
1E248	ꨆ	FINAL M	<u>m</u>	
1E249	ꨇ	SIGN FINAL M	<u>m</u>	Indicates a preceding short vowel.
1E24A	ꨈ	FINAL Y	<u>j</u>	
1E24B	ꨉ	FINAL R	<u>u</u>	This sound is not found in all dialects.
1E24C	ꨊ	FINAL L	<u>l</u>	
1E24D	ꨋ	FINAL V	<u>u</u> / u	
1E24E	ꨌ	FINAL SS	<u>z</u>	
1E24F	ꨍ	FINAL S	<u>h</u>	
1E250	ꨎ	SIGN FINAL H	<u>h</u>	
<b>Digits</b>				
1E251	0	DIGIT ZERO		
1E252	ꨏ	DIGIT ONE		
1E253	ꨐ	DIGIT TWO		
1E254	ꨑ	DIGIT THREE		
1E255	ꨒ	DIGIT FOUR		
1E256	ꨓ	DIGIT FIVE		
1E257	ꨔ	DIGIT SIX		
1E258	ꨕ	DIGIT SEVEN		
1E259	ꨖ	DIGIT EIGHT		
1E25A	ꨗ	DIGIT NINE		

Punctuation & Special Characters				
1E25B	᳚	TANA PATOK SAP		Reduplication Marker
1E25C	᳛	CHAM SHADDAH		Gemination marker. Used by some writers for the word ‘Allah’.
1E25D	᳜	TANA PATOK PHUN		Marks Start of Text
1E25E	᳝	TANA TAMA PHUN		Marks Start of Text
1E25F	᳞	TANA PDAEM PHUN		Marks Start of Text
1E260	᳟	TANA PDAEM IU		Marks Start of Text
1E261	᳠	DANDA		Phrase Divider
1E262	᳡	DOUBLE DANDA		Full Stop
1E263	᳢	TRIPLE DANDA		End of Chapter or Work
1E264	᳣	SIGN TANA TAMAT AYAT		End of Text
1E265	᳤	SIGN TANA TAMAT TAKUE		Poetic Text Divider
1E266	᳥	LUNAR TEN		Represents the Number Ten in the Lunar Calendar
1E267	᳦	SIGN PANGUN		Marks Days of the Waxing Moon
1E268	᳧	SIGN KLAM		Marks Days of the Waning Moon

	1E20	1E21	1E22	1E23	1E24	1E25	1E26
0	 1E200	 1E210	 1E220	 1E230	 1E240	 1E250	 1E260
1	 1E201	 1E211	 1E221	 1E231	 1E241	 1E251	 1E261
2	 1E202	 1E212	 1E222	 1E232	 1E242	 1E252	 1E262
3	 1E203	 1E213	 1E223	 1E233	 1E243	 1E253	 1E263
4	 1E204	 1E214	 1E224	 1E234	 1E244	 1E254	 1E264
5	 1E205	 1E215	 1E225	 1E235	 1E245	 1E255	 1E265
6	 1E206	 1E216	 1E226	 1E236	 1E246	 1E256	 1E266
7	 1E207	 1E217	 1E227	 1E237	 1E247	 1E257	 1E267
8	 1E208	 1E218	 1E228	 1E238	 1E248	 1E258	 1E268
9	 1E209	 1E219	 1E229	 1E239	 1E249	 1E259	
A	 1E20A	 1E21A	 1E22A	 1E23A	 1E24A	 1E25A	
B	 1E20B	 1E21B	 1E22B	 1E23B	 1E24B	 1E25B	
C	 1E20C	 1E21C	 1E22C	 1E23C	 1E24C	 1E25C	
D	 1E20D	 1E21D	 1E22D	 1E23D	 1E24D	 1E25D	
E	 1E20E	 1E21E	 1E22E	 1E23E	 1E24E	 1E25E	
F	 1E20F	 1E21F	 1E22F	 1E23F	 1E24F	 1E25F	

Because the orthography of Western Cham is not standardized, there is some disagreement about the correct usage of the script. Not all characters in this section are used by all writers, who may prefer other characters or diacritics in order to write the same words. Those characters and diacritics are also included in this proposal.

## 2.1 Final Consonants

	Character	Encoding	Name
1	𑜀𑜂𑜆	1E23F	FINAL G
2	𑜀𑜂𑜃𑜫	1E247	FINAL B
3	𑜀𑜂𑜃𑜫	1E248	FINAL M
4	𑜀𑜂𑜃𑜫	1E246	FINAL PH

### 2.1.1 FINAL G (1E23F)

Final G is used in contrast to Final K, with a change of both meaning and pronunciation.

Character	Word	Pronunciation	Meaning	Encoding
FINAL G	𑜀𑜂𑜆	[kək]	coin	1E23F
FINAL K	𑜀𑜂𑜃𑜫	[kəkʔ]	to tie	1E23D

Examples of the use of Final G are shown in [Figures 3 to 6](#).

### 2.1.2 FINAL B (1E247)

Final P can cause diphthongization of a preceding vowel. Final B does not.

Character	Word	Pronunciation	Meaning	Encoding
FINAL P	𑜀𑜂𑜃𑜫	[sauʔ]	smoke	1E245
FINAL B	𑜀𑜂𑜃𑜫	[sap]	sound	1E247

Examples of the use of Final B are shown in [Figures 6 to 9 & 13](#).

### 2.1.3 FINAL M (1E248)

Final M is used when the preceding vowel is long, while the Sign Final M is used when the preceding vowel is short. This means that the word ‘Cham’, for instance, is written with a Sign Final M because the ‘a’ is short. However, the Cambodian Province of Kampong Cham is written using the Final M, because that ‘a’ is long.

Character	Word	Pronunciation	Meaning	Encoding
-----------	------	---------------	---------	----------

FINAL M	ṣṣ	[ɕa:m]	(Kampong) Cham (Province)	1E248
FINAL SIGN M	ṣ̣	[ɕam]	Cham (ethnicity)	1E249

This distinction may also be made with native Cham vocabulary:

Character	Word	Pronunciation	Meaning	Encoding
FINAL M	ṣṣ	[ ɕa:m ]	to meet	1E248
FINAL SIGN M	ṣ̣	[ ɕam ]	to transplant	1E249

Examples of the use of Final M are shown in [Figures 6, 7, 10 & 11](#).

#### 2.1.4 FINAL PH (1E246)

Final PH is used to represent the sound [f] often in words of Arabic origin.

Character	Word	Pronunciation	Meaning	Encoding
FINAL PH	ṣṣṣ	[məʔaf]	pardon	1E246
	ṣṣ	[ʔaf]	to sick (a dog)	

An examples of the use of Final PH is shown in [Figure 12](#).

#### 2.2 FINAL SIGN NG (1E241) vs. SIGN OE (1E233)



Final Sign NG and Sign OE may be identical in some handwriting but they are not considered the same and are also written distinctly in many styles.

These signs often co-occur. In the following example, the word [ciəŋ] meaning ‘to be born’, the lower diacritic is the SIGN OE and the upper diacritic is the FINAL SIGN NG.






Examples of the different shapes and concurrent uses of Sign Final NG (1E241) and Vowel Sign OE (1E233) are shown in [Figures 13, 14 & 16](#).

### 2.3 SIGN L (1E23A) and SIGN SALIH SAP (1E23C)

The term ‘Takay Klak’ is used by many Cham writers to refer to either SIGN L or SIGN SALIH SAP. For some writers these are simply variants of the same character. Some writers only use one of these variants, while some publications make use of both variants without apparent distinction in usage. Other writers do distinguish between the two, reserving SIGN L as a subscript consonant and SIGN SALIH SAP as a diacritic to adjust tenseness. The Cham Language Advisory Committee recommends the use of SIGN L (  ) as a subscript [l] sound in consonant clusters, and the implementation of SIGN SALIH SAP (  ) as a diacritic for the transformation of consonant tenseness when a tense-lax counterpart for that consonant is not available.







It is possible for SIGN L and SIGN SALIH SAP to co-occur.

Characters	Word	IPA	Meaning
MUE		[maj]	grandmother
MUE + SIGN SALIH SAP		[maj̚]	(woman’s name)
MUE + SIGN SALIH SAP + SIGN L		[mlaj]	cotton

Examples of differential usage of these Takay Klaks are found in [Figures 15 & 17-21](#).

### 2.4 Modified Nasal Consonants

There are a total of eight nasal consonant characters representing 4 nasal sounds followed by either the inherent vowel [iʔ] or [aʔ].

Character	Pronunciation	Name	Encoding
	[ŋiʔ]	NGUE	1E204
	[ŋaʔ]	NGA	1E205
	[ɲiʔ]	NHUE	1E20A
	[ɲaʔ]	NHA	1E20B
	[ɲiʔ]	NUE	1E211
	[naʔ]	NA	1E212

𑜋	[miʔ]	MUE	1E21B
𑜌	[maʔ]	MA	1E21C

Some writers use the SIGN SALIH SAP (1E23C) to modify the letters NGUE and NHUE in contexts where other writers would use NGA and NHA. This is not a universal practice. Writers who use NGA and NHA make a distinction between these characters and NGUE and NHUE plus SIGN SALIH SAP.

For these writers SIGN SALIH SAP is used to modify the tenseness of a consonant (tense consonants become lax and lax consonants become tense). In contrast, the letters NGA and NHA are used not to modify the consonant tenseness but to alter the following vowel sound.

The Cham Orthography Guide published in 2011 makes a point that NGA and NHA are not simply NGUE and NHUE with the addition of SIGN SALIH SAP. There are several examples of SIGN SALIH SAP co-occurring with these modified nasal consonants, demonstrating that one is not a substitute for the other.

Examples of the use of NGA, NHA and Nasals Consonants co-occurring with SIGN SALIH SAP are shown in [Figures 21-28](#).

## 2.5 SIGN OEK (1E232)

Proposal SALIH SAP/16-198 includes Sign U. This vowel sign is present in Eastern Cham and is used when transcribing Eastern Cham into Western Cham for the purposes of language learning or comparison. It is also used by speakers of some dialects of Western Cham to reflect their own pronunciation of certain words, where it unambiguously represents the short vowel [iʔ]. The Cham Language Advisory Committee (CLAC) believes this vowel sign should be retained in order to facilitate these activities. However, because the original shape of the vowel sign is similar to SIGN L, the committee has proposed an alternate shape in order to avoid confusion.

[Figure 29](#) shows the CLAC's decision as published in Mukva #11 in 2018.

## 2.6 FINAL SIGN K (1E23E)

FINAL SIGN K is an optional ligature used by some writers as an alternative to FINAL K. It is an abbreviation which sometimes causes the preceding consonant appear to be a final consonant or a consonant accompanied by SIGN SALIH SAP.

Examples are found in [Figure 30](#).

## 2.7 Jibril (1E217)



There exists a special character to represent the [6] sound in the name of the Archangel Jibril (Gabriel).

An example is found in **Figure 31**.

### 3.0 Collation Order

The collation order of Western Cham used by the CLAC is drawn from the traditional presentation order of characters that is common in the Cham community and which would seem sensible and intuitive to Cham speaker literate in Khmer. This order is mostly the same as that found in the Cham-Khmer Dictionary published by the Islamic Community Kan Imam-San of Cambodia under the direction of the late Ong Knur Kay Tam in 2011, although with some differences.

It is clear that the order begins with the consonants KA and KHA and then continues with the remaining consonants in the same order as would be expected by a speaker of Khmer or another language written in an Indic script - ending with the vowel carrier A and then the subsequent independent vowels. The common name for the script is itself Ka-Kha, which, like ABC, is derived from the first letters of the alphabet. For this reason, collation orders starting with vowels or vowel carriers should be rejected.

An example of a typical public display of the Cham script is found in [Figure 2](#).

$\mathcal{N} < \mathcal{N}^{\flat} < \mathcal{O} < \mathcal{V} < \mathcal{V}^{\flat} < \mathcal{V}^{\sharp}$

$\gamma < \gamma^* < \nu^* < \pi < \vartheta < \vartheta^* < \vartheta^*$

$$6\sqrt{2} < 10\sqrt{2} < 17 < 17 < 37 < 37 < 337 < 37$$
$$\mathcal{N} < \mathcal{V} < \mathcal{W} < \mathcal{X} < \mathcal{Y} < \mathcal{Z} < \mathcal{A} < \mathcal{B} < \mathcal{C}$$
$$\mathcal{V} < \mathcal{I} < \mathcal{E} < \mathcal{H} < \mathcal{O} < \mathcal{G} < \mathcal{F}$$

ကု < ဖု < ဖု < ဖု < က < ခ < က < က < ဃ

$$\emptyset \prec \{0\} \prec \{0, 1\} \prec \{0, 1, 2\} \prec \dots$$

$\aleph < \beth < \aleph < \beth < \overset{\circ}{\phantom{0}} < \aleph < \beth < \aleph < \beth < \aleph < \beth < \overset{\circ}{\phantom{0}} < \beth <$

$$m < n < p < q < r < s$$

## 4.0 Cham Dialects and Orthography

Western Cham dialects, including those spoken in southern Vietnam are all mutually intelligible. However, the lack of an orthographic standard has meant that local dialect may influence how individuals choose to write. There are three important divisions in Cham dialects that impact preferred orthographies (Leb Ke 2020).

### 4.0.1 Heavy and Light Accents

Cham dialects or accents are frequently classified by Cham people themselves as being either ‘heavy’ or ‘light’. Heavy accents exhibit aspirated lax consonants [k<sup>h</sup>a:, c<sup>h</sup>a:, t<sup>h</sup>a:, and p<sup>h</sup>a:], which are characterized by a breathy phonation of the following long vowel. Light accents lack these consonants and pronounce words containing them as aspirated tense consonants without breathiness [k<sup>h</sup>aʔ, c<sup>h</sup>aʔ, t<sup>h</sup>aʔ, and p<sup>h</sup>aʔ]. Since aspirated lax consonants are represented by their own characters, speakers of light dialects may not use these letters, or use them inconsistently.

Character	Word	Pronunciation	Meaning	Encoding
GHA	𑜀𑜂𑜆𑜐	[k <sup>h</sup> aʔ]	to stop	1E203
KHA	𑜀𑜂𑜆𑜐	[k <sup>h</sup> aʔ]		1E201
JHA	𑜀𑜂𑜆𑜐	[c <sup>h</sup> a:ʔ]	strict	1E209
CHA	𑜀𑜂𑜆𑜐	[c <sup>h</sup> a:ʔ]		1E207
DHA	𑜀𑜂𑜆𑜐	[t <sup>h</sup> aŋ]	knife	1E210
THA	𑜀𑜂𑜆𑜐	[t <sup>h</sup> aŋ]		1E20E
BHA	𑜀𑜂𑜆𑜐	[p <sup>h</sup> ak]	fermented fish	1E21A
PHA	𑜀𑜂𑜆𑜐	[p <sup>h</sup> ak]		1E218

### 4.0.2 Realization of the Velar Approximant

Some Cham dialects include the velar approximant [ɰ] as a final consonant. This is represented in writing using Final-R. In Cham dialects that do not have this sound, the final vowel is lengthened instead. The effect is similar to the difference between the rhotic and non-rhotic dialects of English. For some speakers the word ‘cake’ is [haɰ], while for others it is [ha:]. Speakers who say [haɰ] will also distinguish the words [paɰaw] ‘to pay back’ and [paɰa:] ‘crocodile’<sup>1</sup>, and write them differently,

<sup>1</sup> Cf. (Malay: bayar ‘to pay’, buaya ‘crocodile’)

while for speakers without this feature, these words are homophones – both pronounced [paja:], and may be written alike.

Character	Word	Pronunciation	Meaning	Encoding
Final R	𑜏𑜤𑜂𑜫	[hau]	cake	1E24B
Sign AA	𑜏𑜤	[ha:]	cake	-
Final R	𑜏𑜤𑜂𑜫𑜏𑜤	[pajau]	to pay back	1E24B
Sign AA	𑜏𑜤𑜂𑜫	[paja:]	crocodile	-

#### 4.0.3 Unstressed Vowel Deletion

Sesquisyllabicity is a feature of all dialects of Cham. This means that the unstressed pre-syllable of a two syllable word tends to be reduced. This reduction can result in the loss of the vowel and the creation of a consonant cluster. This has two common consequences for orthography:

1. When the onset of the pre-syllable is a stop consonant and the onset of the main syllable is [h], then these sounds may fuse causing the word be pronounced with an aspirated rather than unaspirated stop. This means that the word ‘new’ is [pahu] for some speakers, and [p<sup>h</sup>au] for others. Since the Cham script has different characters for aspirated and unaspirated consonants, different speakers may write these words using different initial letters.

2. If the initial consonant of the main syllable is either [l] or [ɣ] then the subscript forms of these consonants are used. This leads to variant forms such as [pulan] and [plan], for ‘month’ and [kaɣae] and [kɣae], for ‘other’. These consonant clusters are realized by the use of SIGN L and SIGN R.

Character	Word	Pronunciation	Meaning	Encoding
PA	𑜏𑜤𑜂𑜫𑜏𑜤	[pahau]	new	1E215
PHA	𑜏𑜤𑜂𑜫𑜏𑜤	[p <sup>h</sup> au]		1E218
(no ligature)	𑜏𑜤𑜂𑜫𑜏𑜤	[pulan]	month	(no ligature)
SIGN L	𑜏𑜤𑜂𑜫𑜏𑜤	[plan]		1E23A
(no ligature)	𑜏𑜤𑜂𑜫𑜏𑜤	[kaɣa:]	other	(no ligature)
SIGN R	𑜏𑜤𑜂𑜫𑜏𑜤	[kɣa:]		1E239

## 5.0 Figures

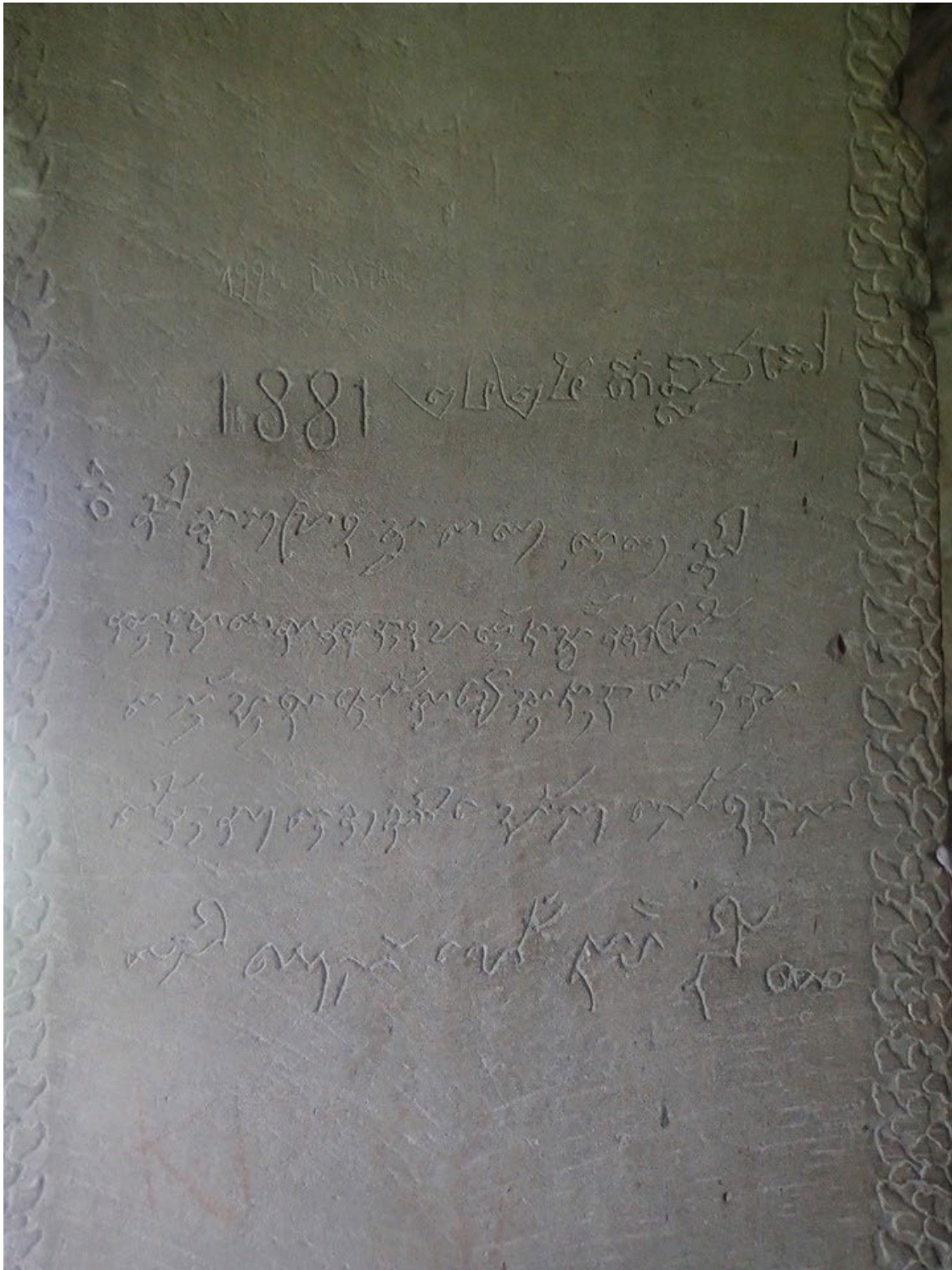
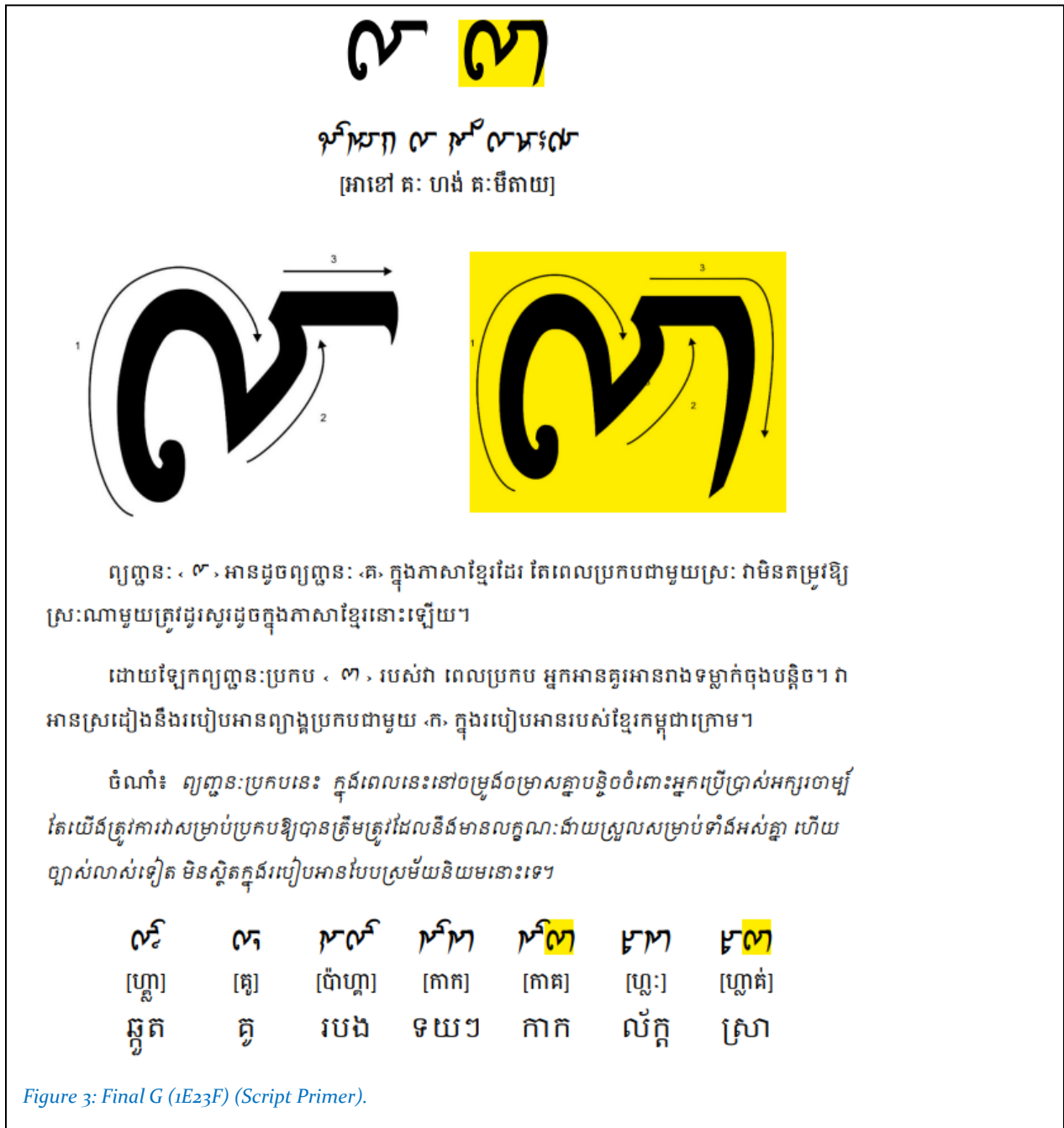


Figure 1: Cham graffiti at Angkor Wat in 1881. Subsequent ethnographic investigation in Kampong Siem, the village of origin of the carver, revealed that the Cham script was in use there until the 1970s. This is evidence for the historic and widespread use of the script beyond the Imam San religious denomination.



Figure 2: Cham Script T-Shirt starting with the letters Ka and Kha, which gives the name of the script – Ka-Kha.





**សេចក្តីត្រូវការ**

ដូចប្រជាជនផ្សេងៗទៀតទាំងអស់ដែលរស់នៅក្នុងប្រទេសនេះដែរ សហគមន៍ចាមបានទទួលខុសត្រូវទៅយ៉ាងខ្លាំង

ក្នុងអំឡុងរបបប៉ុលពត (១៩៧៥-១៩៧៩)។ នៅពេលនោះ ពួកគេត្រូវគេហាយរាល់មិនឲ្យប្រតិបត្តិសាសនារបស់ពួកគេមិនឲ្យនិយាយភាសារបស់ពួកគេ និងមិនឲ្យគោរពប្រពៃណី វប្បធម៌របស់ពួកគេ ដែលនោះជាផ្នែកមួយឲ្យពួកគេមានលក្ខណៈខុសប្លែកពីប្រជាជនខ្មែរ។ សៀវភៅនិងអត្ថបទសាសនាត្រូវបំផ្លិចបំផ្លាញបាត់បង់ក្នុងអំឡុងពេលនោះ ហើយពិតជាមានគ្រោះថ្នាក់យ៉ាងខ្លាំងក្នុងការបង្រៀនភាសា ឬក៏ត្រូវគេរកឃើញថាមានកម្មសិទ្ធិនៃរាងកាយចាមណាមួយ។

**Need for a literacy program**

Like all other people living in the country, the Cham community suffered greatly during the Pol Pot regime (1975-1979). During this time, they were prohibited from practicing their religion, speaking their language and observing any cultural traditions that might distinguish them from Khmers. Religious books and manuscripts were destroyed or otherwise lost during this time, and it would have been very dangerous to either teach the language or be found in possession of any Cham documents.

*Figure 4: Final G (Ie23F) (Rediscovering Cham Heritage).*

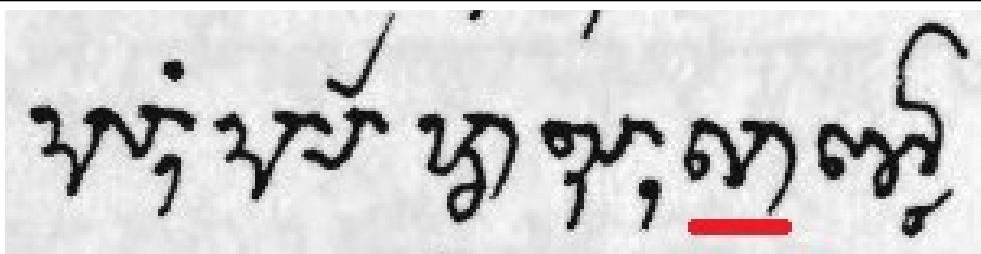


Figure 5: Final G (1E23F) This example is taken from page 5 of 'Songs to Bind Three Generations' by Hammat Yousoh, and shows the use of Final G. Here (Second to last character) it is used to write the word Oknha a borrowing from Khmer (ឧកញ៉ា).

The handwritten form is: 

# ព្រឹត្តិបត្រមុន្នា នាពេលឆ្នាំ

ឆ្នាំទី ១០

ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១ ឆ្នាំទី ០១ ឆ្នាំទី ០១

ឆ្នាំទី ០១ ឆ្នាំទី ០១ ឆ្នាំទី ០១

ឆ្នាំទី ០១

ឆ្នាំទី ០១ (ឆ្នាំទី ០១) ឆ្នាំទី ០១ ឆ្នាំទី ០១

ឆ្នាំទី ០១ ឆ្នាំទី ០១ ឆ្នាំទី ០១

Figure 6: Final G (1E23F) in red, Final B (1E247) in yellow and Final M (1E248) in blue. (Mukva #10)



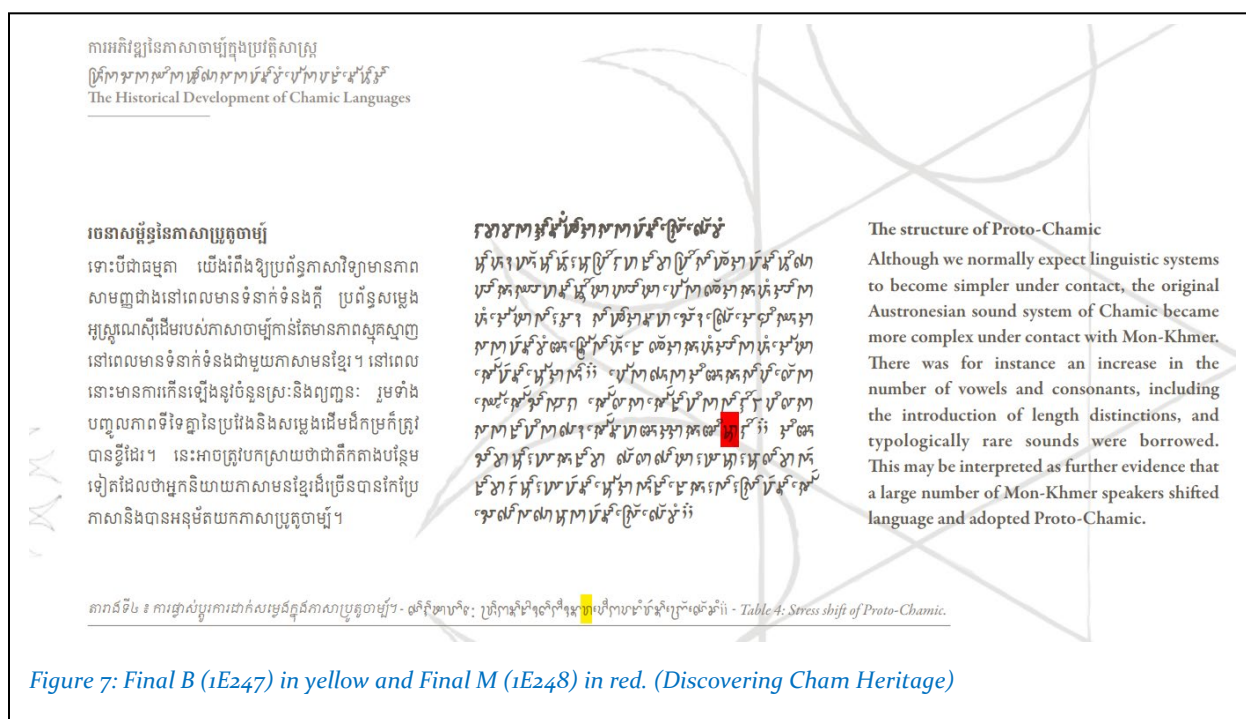


Figure 7: Final B (1E247) in yellow and Final M (1E248) in red. (Discovering Cham Heritage)

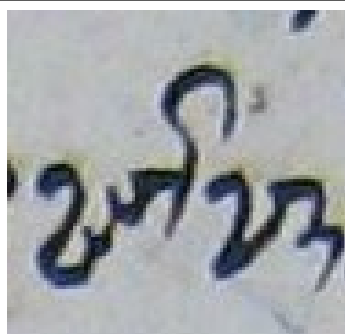


Figure 8: Final B (1E247) The scribe's writing style creates some ambiguity as this word initially appears to be written with the symbol Takay Kak. This may be an effect of his handwriting or an indication that this writer does not use Final B. However, usage of Takay Kak here would mean the word is pronounced [pha:paʔ], but this is not a word. Instead it is pronounced [pha:pʔ], meaning 'people'—with a Final B. In epigraphy, these ambiguities must be resolved on a case by case basis by the researcher studying the manuscript. (Buraq Manuscript)

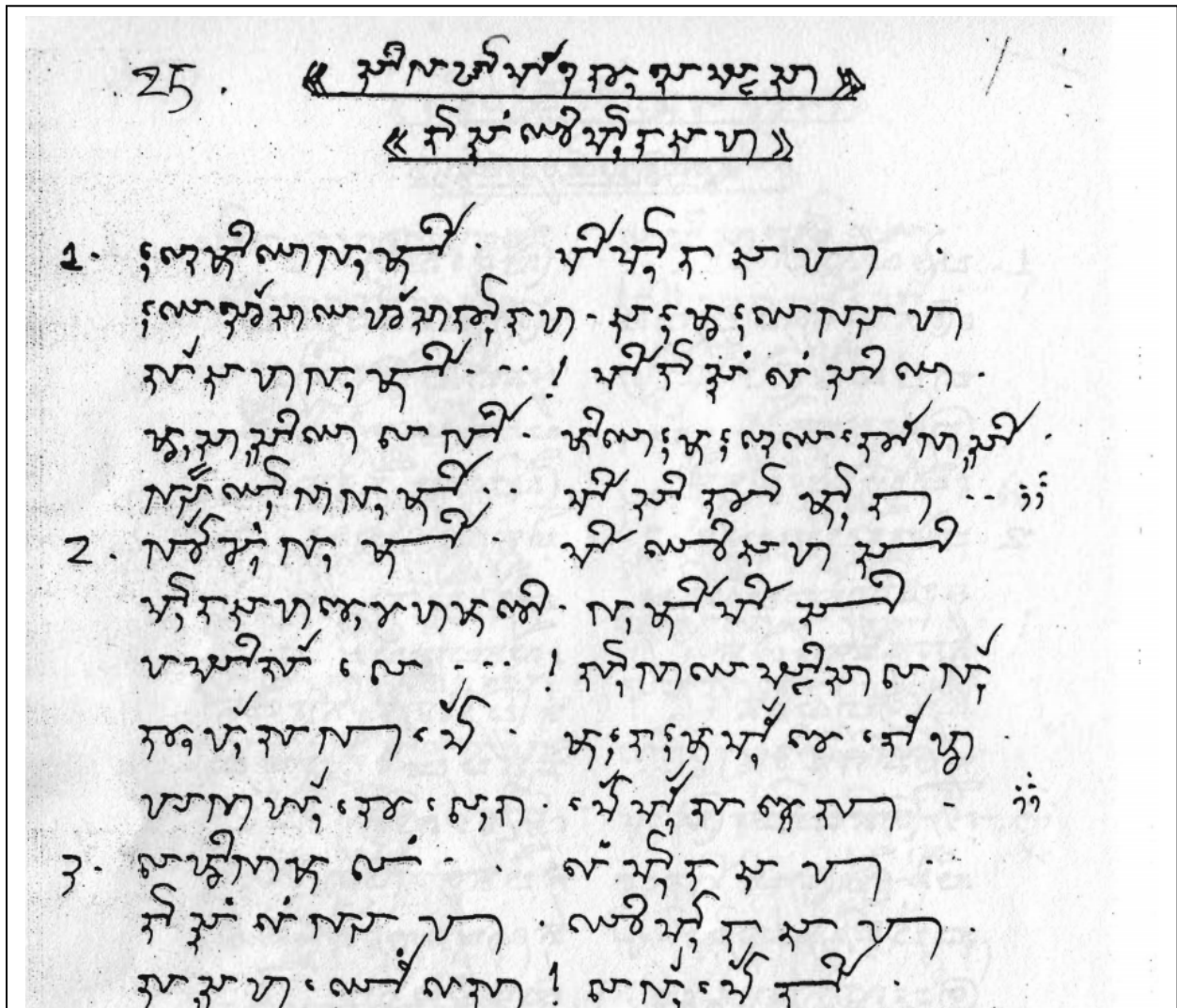
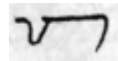
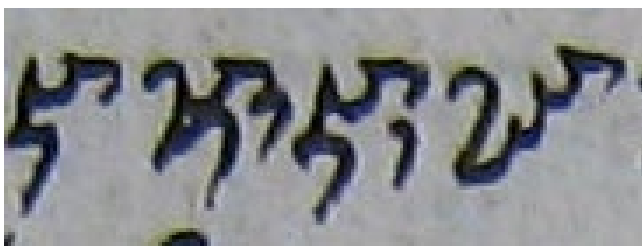


Figure 9: Final B (1E247). This example is taken from 'The Four Meanings of Marriage' by Hammat Yousos.

Final B appears several times on page 25. Handwritten it has the form:





*Figure 10: Final M (1E248) The scribe's writing style creates some ambiguity as this word initially appears to be written with the symbol Takay Kak. However, this would mean the word is pronounced [sami?supha:], but this is not meaningful in this context. This word is pronounced [samsupha:], meaning 'foundation' –with a Final M in medial position. (Buraq Manuscript)*

[illegible]

ကုမ္ပဏီ: ကုမ္ပဏီအသစ်တစ်ခုကို ဖွဲ့စည်းခဲ့သည်။

[illegible]

Figure 11: Final  $M$  (1E248). (Mukva #12)

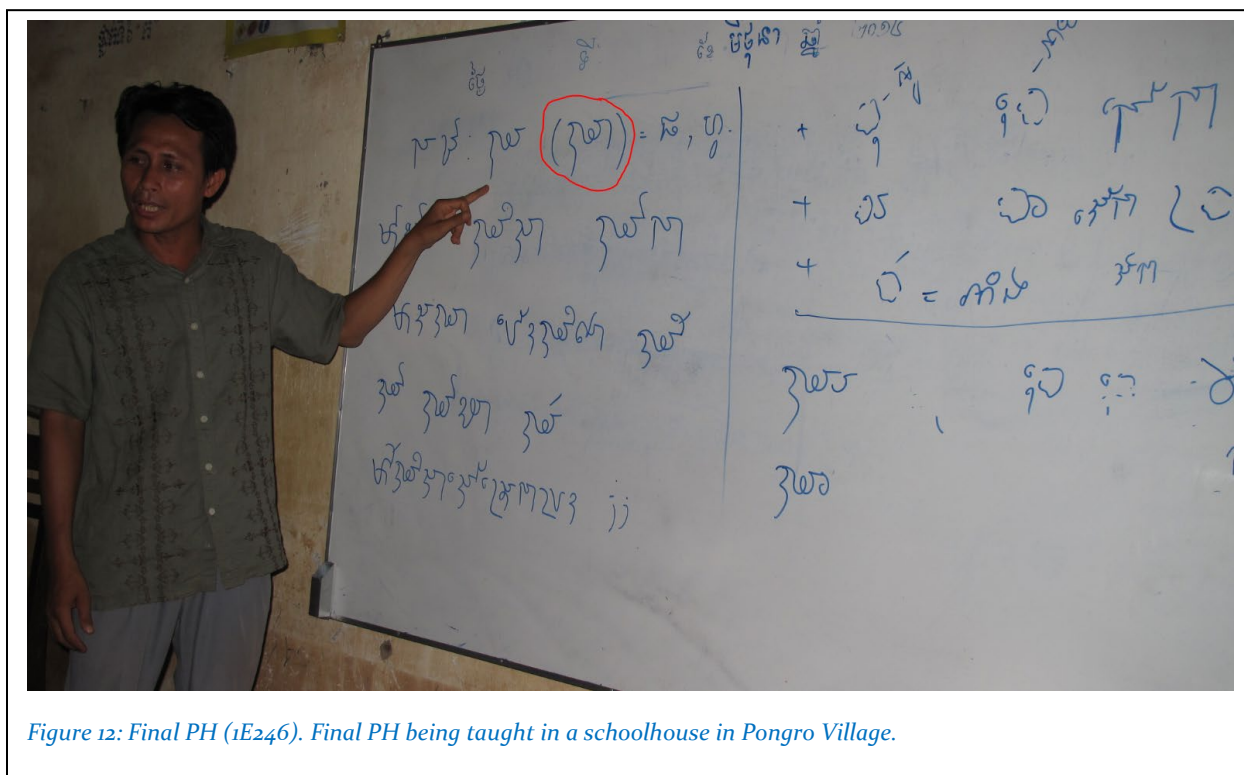


Figure 12: Final PH (1E246). Final PH being taught in a schoolhouse in Pongro Village.



Ἰῆς Ἰῆς Ἰῆς, ἦ

၁. အောက်ဖော်ပြပါများ ..... နှစ်
၂. ၁၉၈၁ ခုနှစ် ..... နှစ်
၃. ၁၉၈၁ ခုနှစ် ..... နှစ်

ᠨᠢᠨᠠᠨᠠᠨᠠᠨ

[illegible][illegible][illegible]

ကံကုဏ္ဍိယ: ဟိုလီ

Figure 13 - Co-occurrence of Sign Final NG (1E241) and Vowel Sign OE (1E233) in green and Final B (1E247) in yellow. (Mukva #9)

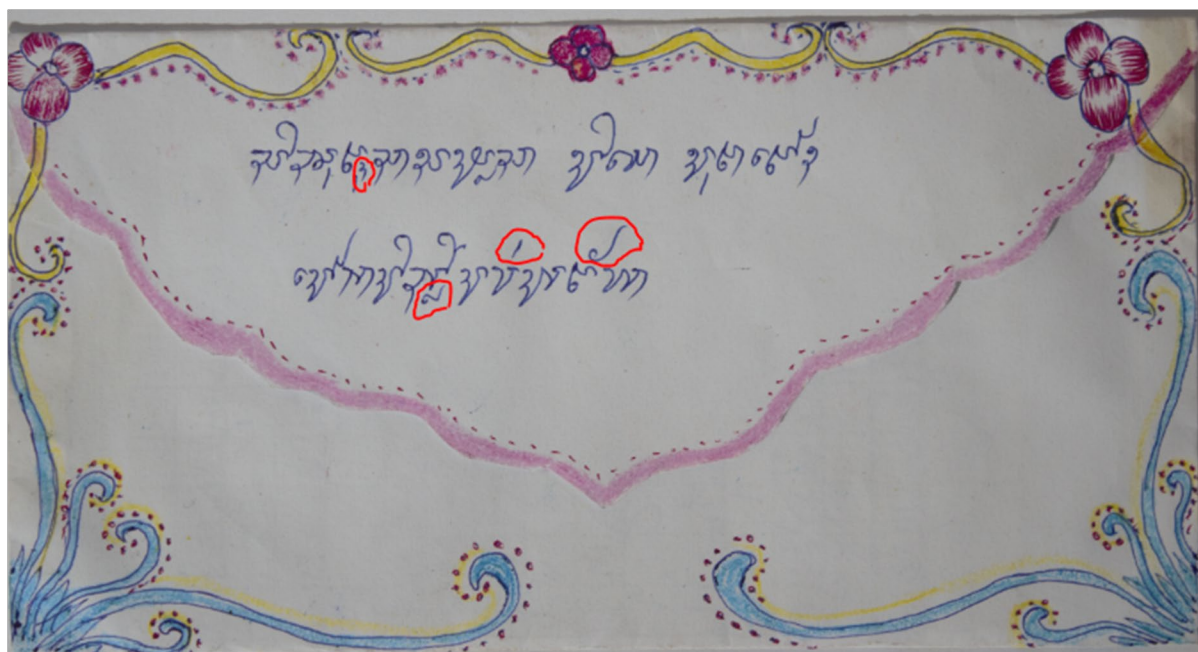
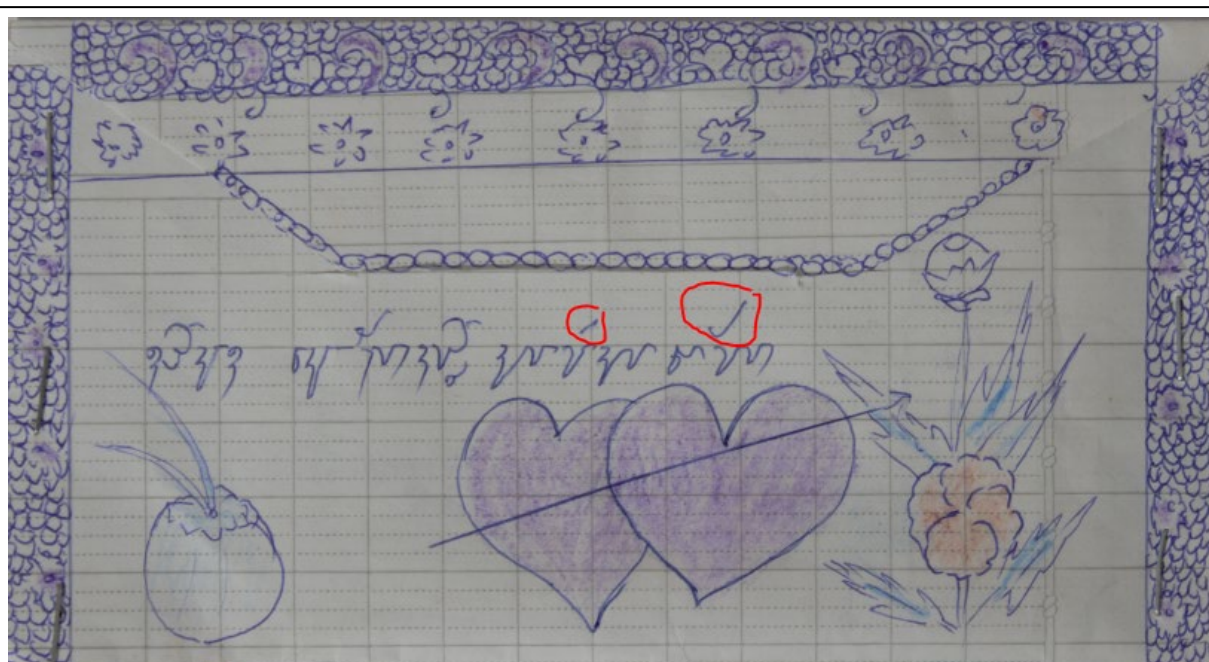


• ខ្មែរព្យាបាល:

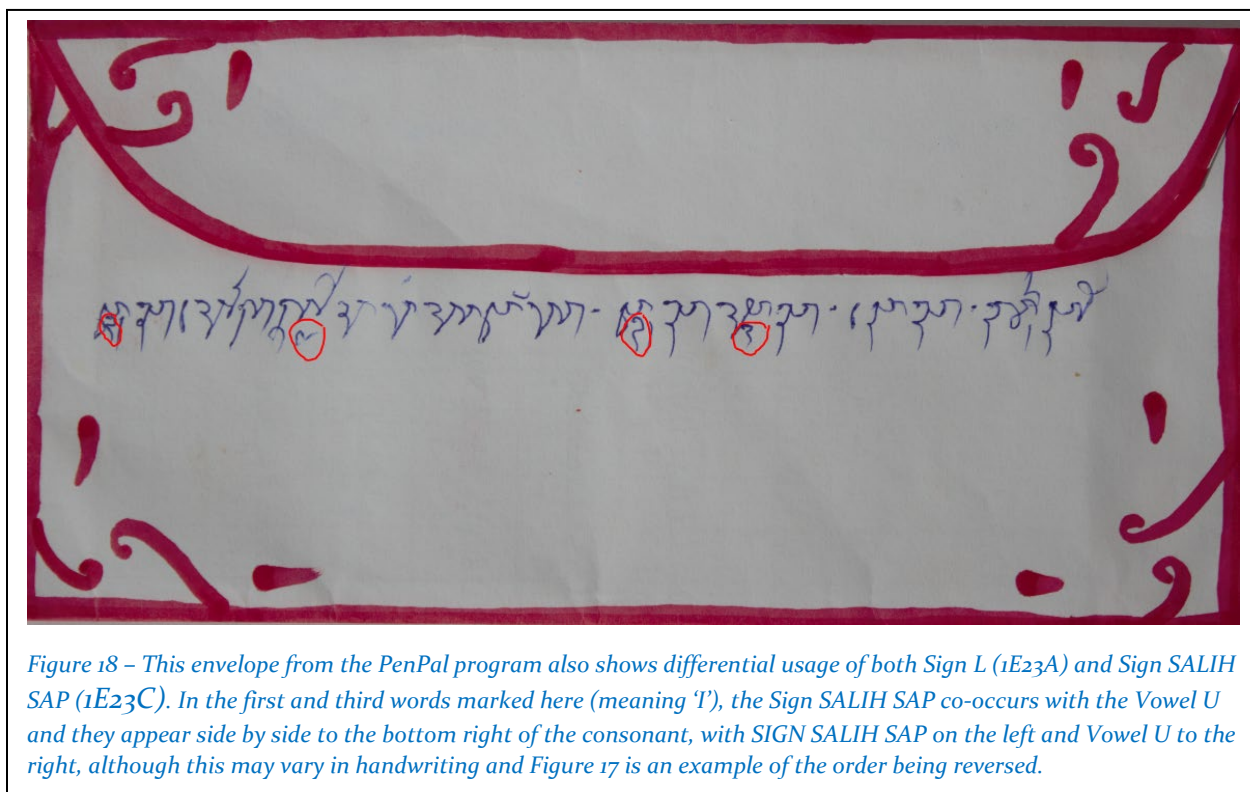
តាកាយ	ជើងព្យាបាល:	ការប្រកប	
៧	ល្អ	៧, ៧	ល្អ , ល្អ
៧	ប្រ	ប្រ	ប្រ
៧	ត្រ	ត្រ	ក៏
៧	យ្យ	យ្យ	ក៏
៧	ក្រ	ក្រ, ក្រ	កាក, ឡាក

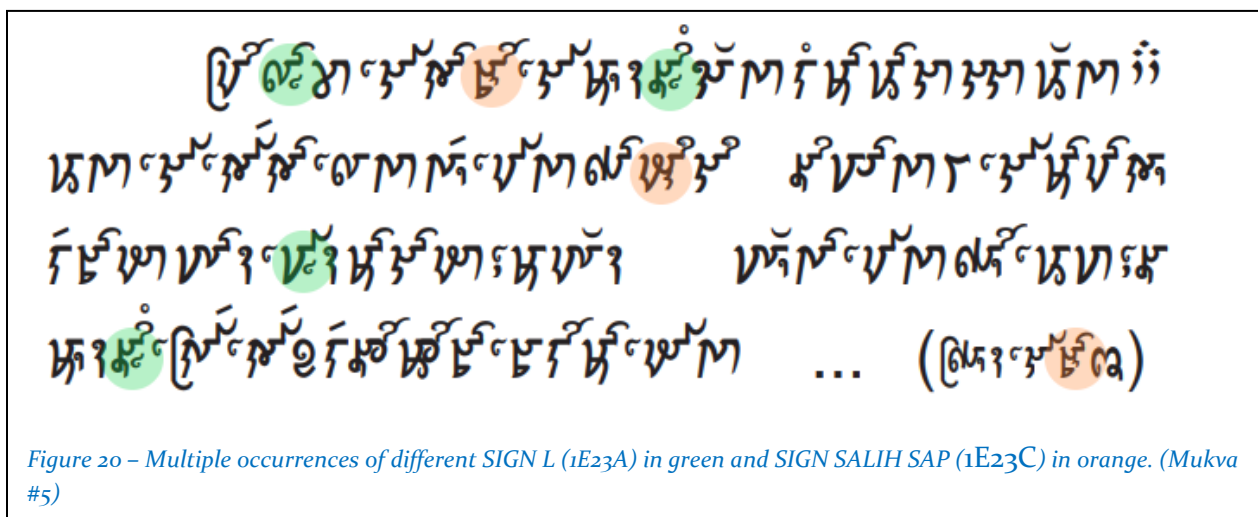
Figure 15 – Page iv of the Cham –Khmer Dictionary published by the Krom Kan Imam-San Community showing differential usage of symbols, both of which are frequently referred to as Takay Klak. In this text, it is suggested that the second Takay Klak be used as a final glottal stop (Takay Kak) rather than as a method for adjusting tenseness. An examination of the dictionary entries reveals that this distinction is not maintained in any systematic way. Nevertheless, this establishes that these two forms of Takay Klak are not merely stylistic variants for all authors. In this proposal they appear as SIGN L and SIGN SALIH SAP.





Figures 16 & 17 - Co-occurrence of Sign Final NG (1E241) and Vowel Sign OE (1E233) above the consonants. The second image includes two forms of SIGN SALIH SAP (1E23C) below the consonants – modifying consonant tenseness and the second SIGN L (1E23A) acting as a subscript consonant. These are envelopes from a PenPal program that encouraged Cham children to write to children in other villages using their language and script as a way of teaching the value and utility of mother-tongue literacy.





- ✚ ព្យញ្ជនៈ [២] នេះមានឈ្មោះថា <ង្ហាត់> ស្រដៀង<ង>ដែរតែមានស្រះកប់ត្ប[័]។
- ✚ ព្យញ្ជនៈ [៣] នេះមានឈ្មោះថា <ង្ហាត់> ស្រដៀង<ង>ដែរតែមានស្រះកប់ត្ប[័]។
- ✚ ព្យញ្ជនៈ [៤] នេះមានឈ្មោះថា <ញាត់> ស្រដៀង<ង>ដែរតែមានស្រះកប់ត្ប[័]។
- ✚ ព្យញ្ជនៈ [៥] នេះមានឈ្មោះថា <ង្ហាត់> ស្រដៀង<ង>ដែរតែមានស្រះកប់ត្ប[័]។

ចំណាំ! ព្យញ្ជនៈទាំងនេះមិនមែនបានមកពី ៖

២ ≠ ២ + ័    ៣ ≠ ៣ + ័    ៤ ≠ ៤ + ័    ៥ ≠ ៥ + ័

Figure 21 – Comparison of Nasals Consonants with vowel UE and modified Nasal Consonants with vowel A. The bottom line indicates that the Takay Klak (either SIGN L or SIGN SALIH SAP) is not a substitute for the use of the modified nasal. Character NGA (1E20B) (Cham Orthography Guide).

តារាងទី៣ ៖ ពាក្យចាមដែលមានភាពត្រូវគ្នាផ្នែកសម្លេងមួយចំនួនៗ - ព័ត៌មានលម្អិត: តាមការស្រាវជ្រាវរបស់លោកស្រី ឈីន ឈីន - Table 3: Some corresponding Chamic words.

ម៉ាឡេ	ប្រូតូចាម	អាសេ	ច្រូ	កំភ្លែងខាងជើង	ចាមខាងកើតផ្លូវការ	
អ៊ីដ្រ	ឆ្រីដ្រ	ឆ្រីដ្រ	ឆ្រីដ្រ	ឆ្រីដ្រ	ឆ្រីដ្រ	
Malay	Proto-Chamic	Acehnese	Chru	Northern Raglai	Formal Eastern Cham	
darah	*darah	darah	drah	darah	ṭarāh	‘ឈាម’ ឆ្រីដ្រ ‘blood’
dara	*dara	dara	dra	dara	ṭara	‘ក្រមុំ’ ឆ្រី ‘girl’
bulu	*bulow	bulew	blow	bilow	pilow	‘ពាម’ ឆ្រី ‘hair’
bulan	*bula:n	buluon	ea bla:n	ia bila:t	pilan	‘ព្រះច័ន្ទ’ ឆ្រី ‘moon’
baharu	*bahrow	baro	bərhəu	bahrəu	pirəw	‘ថ្មី’ ឆ្រី ‘new’
tahun	*thun	thon	thun	thut	thün	‘ឆ្នាំ’ ឆ្រី ‘year’
tangan	*tanga:n	---	təŋa:n	taŋän	taŋin	‘ដៃ’ ឆ្រី ‘hand’

Figure 22 - Character NGUE (1E204) with SIGN SALIH SAP (1E23C). (Rediscovering Cham Heritage)



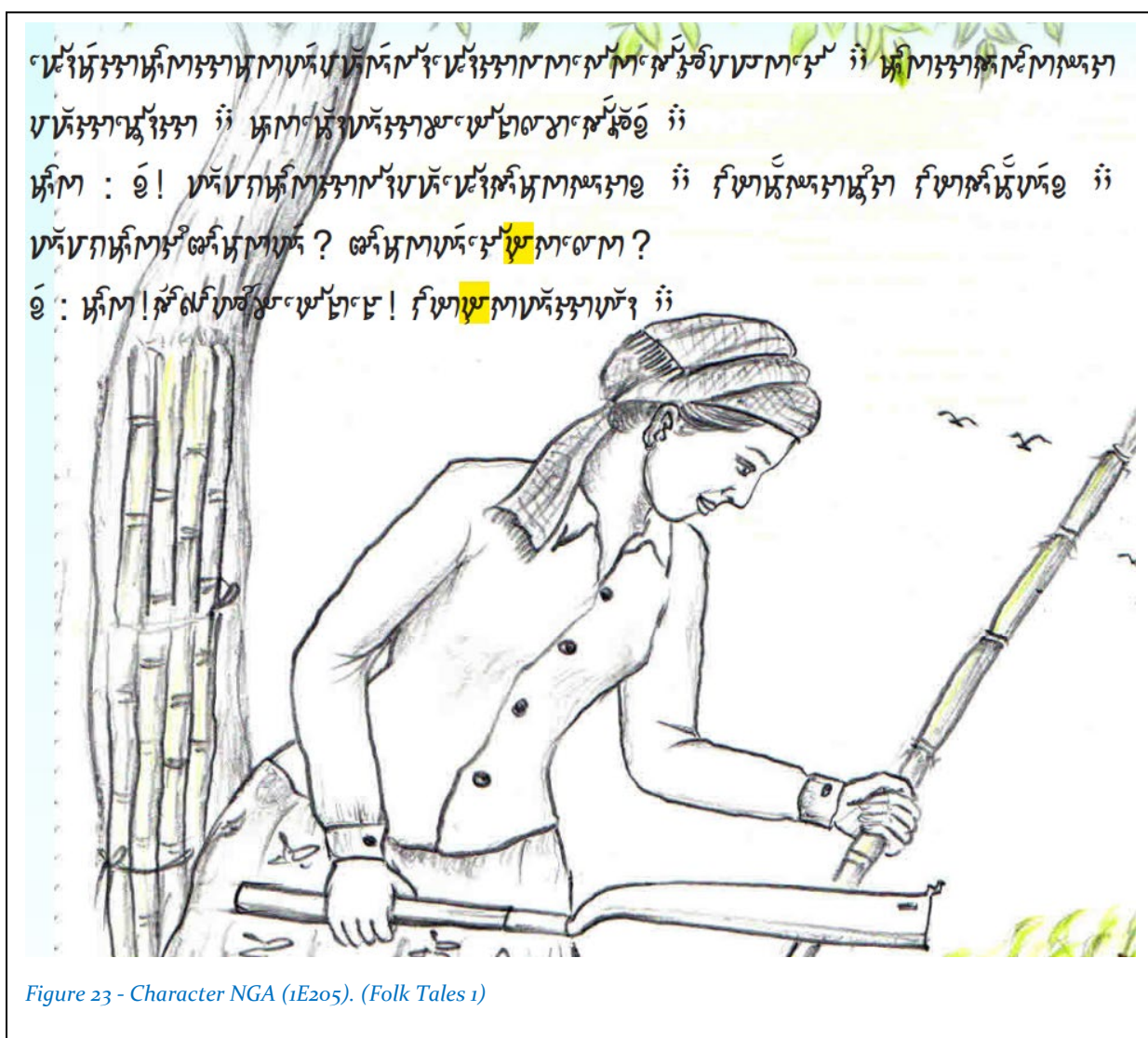


Figure 23 - Character NGA (1E205). (Folk Tales 1)

The figure shows a page from a document with text in Khmer and Cham scripts. The Khmer text on the left discusses the historical context of Cham language publications in Cambodia, mentioning 'Champa Sa' and 'Ysa Osman'. The Cham text on the right is a historical document snippet, with a yellow box highlighting a specific character 'NHA' (1E20B) and noting it is a 'lax' consonant. The caption below the image identifies this character and its historical context.

[illegible][illegible]





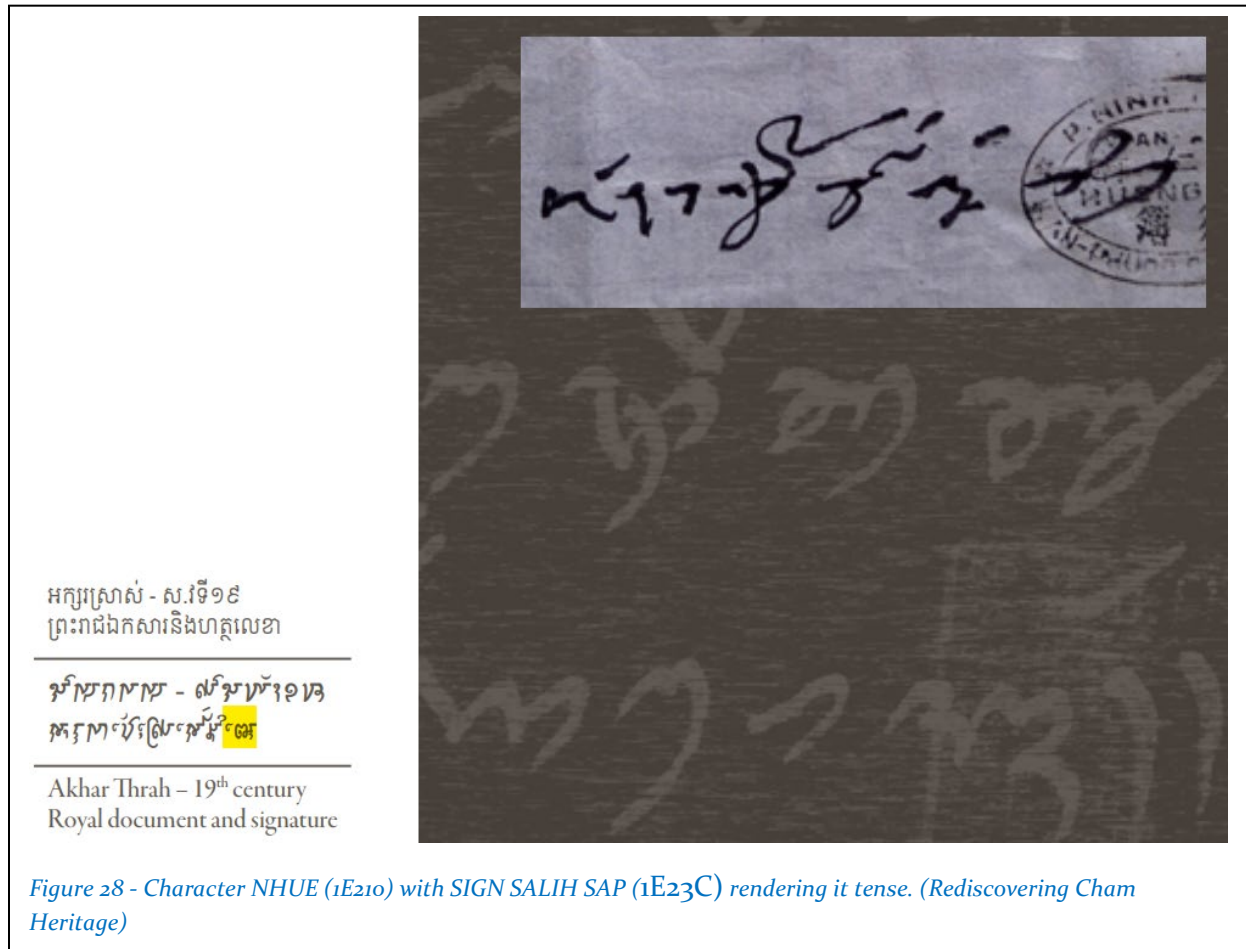
១៣. វត្តភ្នំបាទ្យស្រីវិហារភ្នំព្រៃវែង-វត្តភ្នំបាទ្យស្រីវិហារ ២០១៩

(ស្រីម្នីម្នី) ... ស្រីម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី  
 ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី  
 ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី  
 ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី

៣ ម្នីម្នី ម្នីម្នី ម្នីម្នី

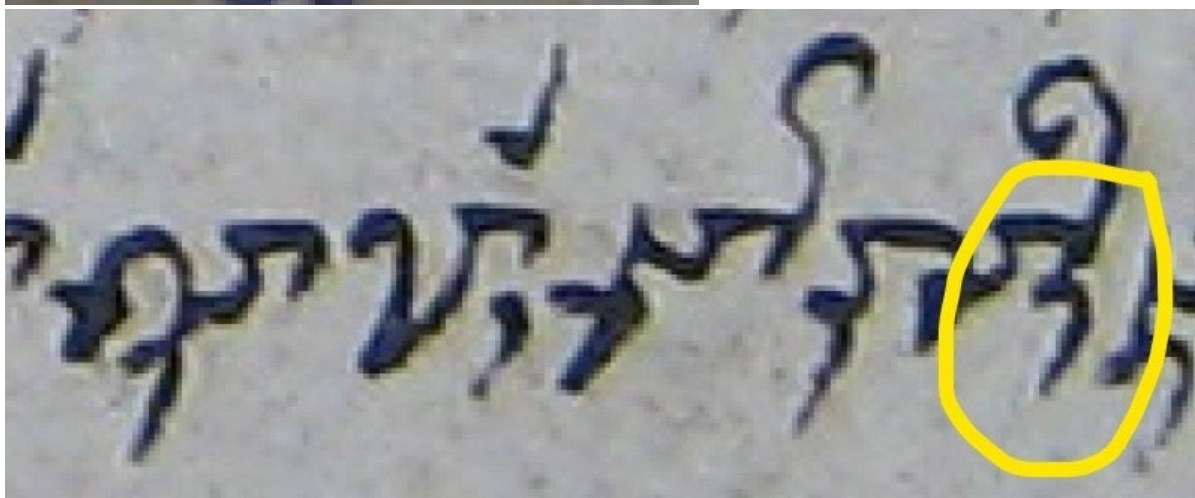
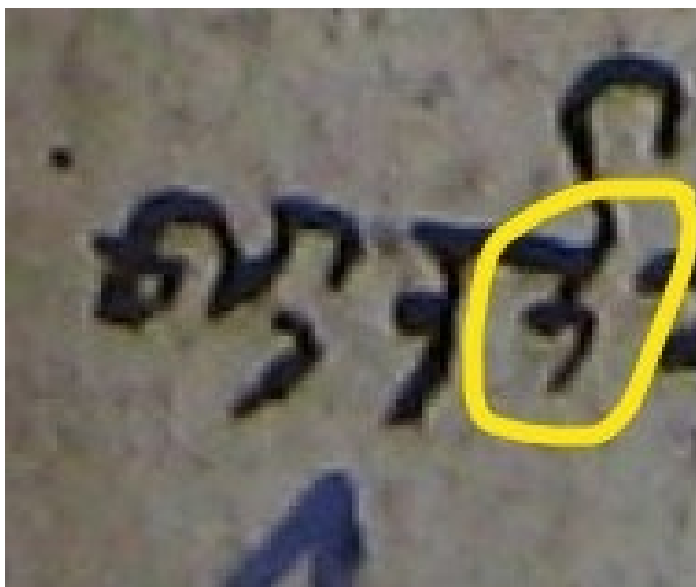
ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី  
 ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី  
 ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី ម្នីម្នី

Figure 27 - Character NUE (1E211) with SALIH SAP (1E23C) in yellow and Character NHUE (1E20A) (tense) with SIGN SALIH SAP in red. (Mukva #5)









*Figure 30 – FINAL SIGN K. The first word [kuyaḱ̄] meaning ‘leader’ and the second word being a personal name [ʔapulaḱ̄iḱ̄]. (Buraq Manuscript)*

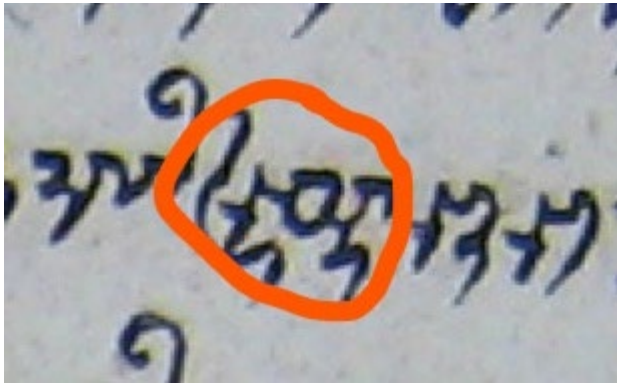


Figure 31 – Ligature representing the name of the Archangel Jibril (Buraq Manuscript).

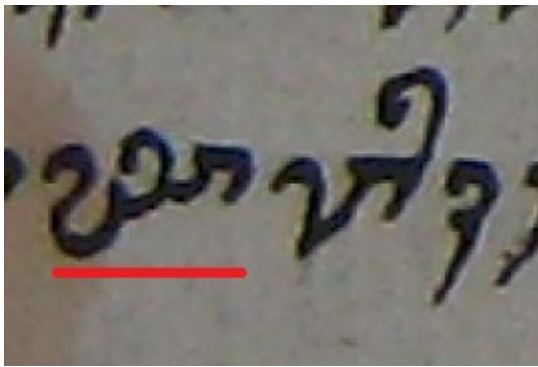


Figure 32 – Character A DHA being used to write the word [ʔapɪh] meaning ‘all’. (Buraq Manuscript)

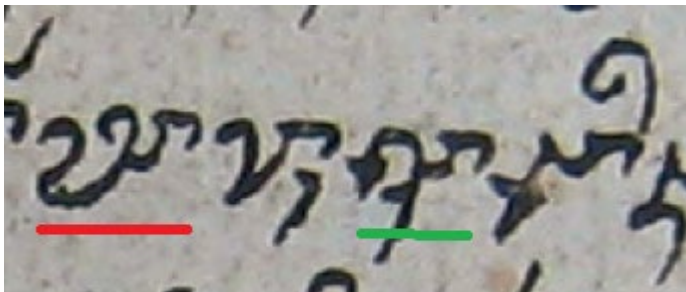
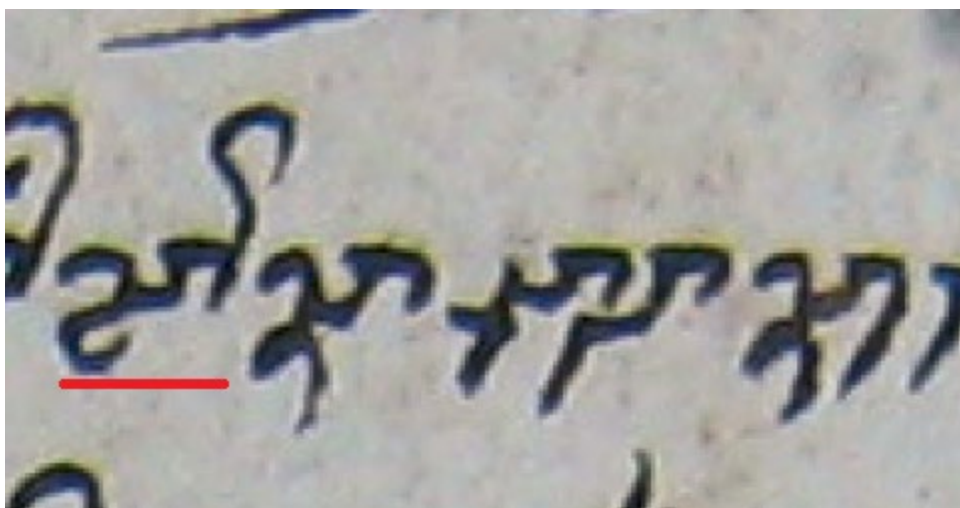
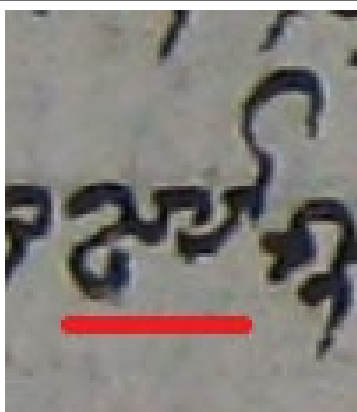


Figure 33 – Character A BA being used to write the first part of the proper name [ʔapuʔali]. The second part of the name [ʔali] is written with the more conventional Character A. (Buraq Manuscript)



*Figure 34 – Character A U being used to write the word [ʔa:nɨkʰan] meaning ‘story or account’. (Buraq Manuscript)*



*Figure 35 – Character A U being used to write the word [ʔɛə] meaning ‘water’. (Buraq Manuscript)*

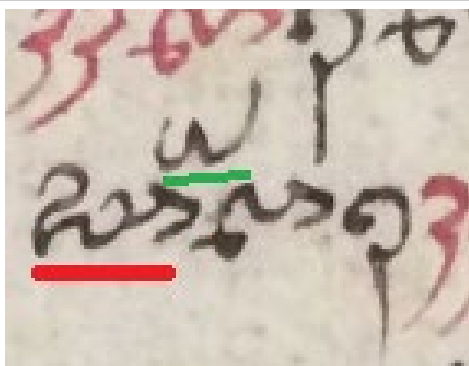


Figure 36 – Character A DHA (underlined in red) being used in conjunction with CHAM SHADDAH (underlined in green) to write the word ‘Allah’ with a doubled [l] sound. (Unnamed Manuscript)

## 6.0 References

Brunelle, Marc. (2005). *Register in Eastern Cham: Phonological, Phonetic and Sociolinguistic Approaches* (PhD diss). Cornell University.

Brunelle, Marc and Graham Thurgood. (2015). The Historical Development of the Chamic Languages in “Rediscovering Cham Heritage in Cambodia: Language, Script, and Community”. First Edition ed. Phnom Penh: Naga Editions.

Cœdès, Georges. (1939). La plus ancienne inscription en langue cham (Inscription rupestre de Đông Yên Châu, province de Quảng Nam, Annam) in “New Indian Antiquary: A volume of Eastern and Indian Studies in honour of F.W. Thomas, C.I.E., edited by S. M. Katre and P. K. Gode. Bombay: Karnatak Publishing House

Hammat Yousoh. (2014). សៀវភៅកំណាព្យ ភរិយាស្វាមីនីយប្បនប្រការណ៍ (The Four Meanings of Marriage). Self-Published. Pursat, Cambodia.

———. (2004). កំណាព្យរៀបរាប់ ជាប់បីជំនាន់ (Songs to Bind Three Generations). Self-Published. Pursat, Cambodia.

Islamic Community Kan Imam-San of Cambodia. (2011). Cham –Khmer Dictionary. Sreybrey Village, Kampong Chhnang Province.

Leb Ke. (2018). The Presence of Cham Inscriptions at Angkor Wat. *Kambuja Soriya*, 65(3), 79-98.

Leb Ke. (2020). *A Phonetic and Phonological Description of Western Cham, Ou Ruessei Commune, Kampong Chhnang Province* (MA Thesis). Royal University of Phnom Penh.

López-Cortina, Jorge, and Alberto Pérez-Pereiro, eds. (2015). *Rediscovering Cham Heritage in Cambodia: Language, Script, and Community*. First Edition ed. Phnom Penh: Naga Editions.

Pérez-Pereiro, Alberto, and Leb Ke. (2014). *Cham Script Primer*. Phnom Penh: Naga Editions.

Pérez-Pereiro, Alberto, and Leb Ke, eds. (2014). *Cham Folktales 2*. Phnom Penh: Emerging Markets Consulting.

———. (2013). *Cham Folktales*. Phnom Penh: Emerging Markets Consulting.

Pérez-Pereiro, Alberto, and Leb Ke. (2011). *Cham Orthography Guide*. Phnom Penh: Emerging Markets Consulting.

———. (2011). *Cham Literacy Textbook 2*. Phnom Penh: Emerging Markets Consulting.

———. (2011). *Cham Literacy Textbook 1*. Phnom Penh: Emerging Markets Consulting.

Ueki, Kaori. (2011). *Prosody and Intonation of Western Cham* (PhD diss). University of Hawai‘i at Mānoa.

## 6.1 Undated Manuscripts

*Buraq*

*Adab Safi*