Preserving Cham Font through Online Conversion Application

Van Ngoc Sang[1] & Mohamad Bin Bilal Ali[1]

[1] Faculty of Education, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Correspondence: Mohamad Bin Bilal Ali, Faculty of Education, Universiti Teknologi Malaysia, Johor Bahru, Malaysia. E-mail: mba@utm.my

Abstract

The Cham people who are now the minority ethnic in Vietnam speaks with the language familiar used as others Malay groups but differ in their written language. Cham Script inscriptions appear on *Dong Yen Chau* stone stele (Tra Kieu) in 4th century and the Cham are using this script system until today. Ensuring the preservation of Cham language, this study intended to design a tool to convert the EFEO Cham Latin in Malay system to *Cham Akhar Thrah*. The method used is by converting Latin EFEO into intermediate characters code followed by assigning it to *Akhar Thrah* backwards. Cham font conversion application has been created, which has carried out a number of technical requirements, and content conversion ensures correct in vocabulary, semantics and grammar. In this experiment we have checked the accuracy percentage of three Cham poems and results *Ariya Cam Bini* 100% (n=1823); *Ariya Gleng Anak* 99.88% (n=2459); *Nai Mai Mang Makah* 100% (n=2523). Cham font conversion is necessary and meaningful in conservation of Cham script. It will be used in schools, institutions in the country and overseas as well as assist in teaching and learning Cham language.

Keywords: Cham language, Cham script, Cham conversion, Cham font conversion

1. Introduction and Background

The Cham script was a systems of signs derived from the Sanskrit alphabet in southern India, in the long term approach culture of Hindu civilization. The Cham people have used Indian Sanskrit and then they have improved this Sanskrit completely and become the Cham script (Sastri, 1935).

The use of information technology to preserve the endangered language is not new in the world today when digital technology has developed well. Information technology has significantly contributed in the indigenous language education, restoring and preserving of ancient writing system (Grenoble & Whaley, 2006). In Vietnam, some organizations and individuals have studies on designing a number of fonts such as *champa.ttf*, *camtanran.ttf*, *bingudi tanran.ttf*, *cam thrah.ttf*, *Cham Thrah1.ttf* to support for compilation of Cham documents on the computer. However, the opportunity for studying of *Akhar Thrah* for community is limited. Furthermore, identification the typeface of *Akhar Thrah* characters is more difficult than read the Latin typeface. Hence, the EFEO Cham Latin transliteration followed Malay system was established in 1996 to support for reading and writing Cham *Akhar Thrah*.

Font conversion has been widely applied by many researchers to convert different font formats for different purposes. For this study, we propose a Cham font conversion method, which is based on the Cham Latin EFEO to Cham *Akhar Thrah*. In section 1.2, we used methods numbered and assigned to each Cham Latin character corresponds to the Cham *Akhar Thrah*. In section 2.3 (conversion method), Cham Latin EFEO was first converted into intermediate characters code and subsequently converted into the correspondence to the Cham *Akhar Thrah*. In this section we have shown in detail how to convert for each Cham character group.

In Vietnam, there is currently no tool to support for converting Cham Latin to Cham script. Therefore, in this study we carry out a method and construction the Cham font conversion.

1.1 Exploring the Preferable Cham Script for Preservation

In order to explore the preferable Cham script to be preserved, the Figure.1 describes the development process of Cham script and choosing Cham script and Cham Latin to preserve. Based on the result of survey, we develop an application for Cham font conversion. Research activities phases are presented in the diagram as below.
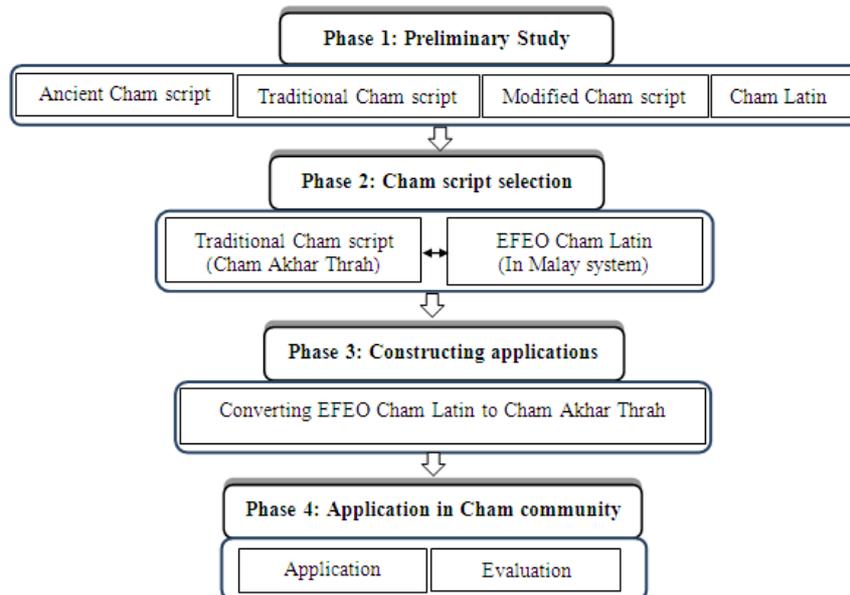
Figure 1. Choosing Cham script to develop applications

There are four phases were taken to develop the application product to preserve Cham language.

Phase 1 Preliminary Study: Firstly, we will synthesize materials related to Cham script, study and systematics all the basic contents. Currently Cham people are using 3 types of script such as Cham script derive from Sanskrit, Jawi script and the Latin script. In this study, we mention only script originated from Sanskrit and Latin. The Cham script has derived from Sanskrit are divided into 3 styles which are ancient Cham script, traditional Cham script and modified Cham script. Latin scripts are including EFEO Cham Latin (based on Malay system and traditional Cham script), and CTCC Latin (based on modified Cham script).

Phase 2 Cham script selection: In order to select the Cham script to be preserved, we will conduct a survey on the most preferable Cham script among the Cham community. The main group who involve in the survey including religious groups, students and Cham intellectuals. Currently in Cham communities, they was selected the script based on the traditional Cham script from the 17th century to preserve, because this script is stable and is stored in multiple documents. For the Latin script, we have decided to use the Latin of Malay system by EFEO applies for the Cham language.

Phase 3 Constructing applications: To preserve Cham script effectively, we carry out solutions using technology applications in preservation Cham language. An interesting application is converting EFEO Cham Latin to Cham script (Akhar Thrah).

Phase 4 Application in Cham community: First of all, this application should be assessed through a forum, to verify the feasibility and completeness of an application program before bring it widely used into the community.

1.2 Assigning Cham Chacracter

Traditional Cham script is a script derived from the Sanskrit which is difference Latin characters. In order to convert the Cham Latin to Cham *Akhar Thrah*, we need to assign values between Akhar Thrah script and Cham Latin respectively as in Table 1.

Table 1. Assigned values between Cham Latin EFEO and Cham Akhar Thrah

Vowels:

Cham Latin EFEO	a	i	u	é	ai	o
Akhar Thrah	ꨀ	ꨁ	ꨂ	ꨄ	ꨀꨰ	ꨀꨯ
Intermediate Character Code (ICC)	0	1	2	3	4	5

Consonants:

Latin EFEO	k	Kh	g	gh	ng	Ng	c
Akhar Thrah	ꩀ	ꩁ	ꩂ	ꩃ	ꩄ	ꩅ	ꩆ
ICC	6	7	8	9	10	11	12
	ch	J	jh	ny	Ny	nj	t
	ꩇ	ꩈ	ꩉ	ꩊ	ꩋ	ꩌ	ꩍ
	13	14	15	16	17	18	19
	th	D	dh	n	N	nd	p
	꩎	꩏	꩐	꩑	꩒	꩓	꩔
	20	21	22	23	24	25	26
	ph	B	bh	m	M	mb	y
	꩕	꩖	꩗	꩘	꩙	꩚	꩛
	27	28	29	30	31	32	33
	r	l	w	s	h	P	S
	꩜	꩝	꩞	꩟	ꩠ	ꩡ	ꩢ
	34	35	36	37	38	39	40

Final Consonant

Latin EFEO	k	ng	c	t	n	p	y	r	l	w	s	h	m
Akhar Thrah	ꩀ	ꩄ	ꩆ	ꩍ	꩑	꩔	꩛	꩜	꩝	꩞	꩟	ꩠ	·
ICC	41	42	43	44	45	46	47	48	49	50	51	52	53

Number:

Latin EFEO	1	2	3	4	5	6	7	8	9	0
Akhar Thrah	ꩢ	ꩣ	ꩤ	ꩥ	ꩦ	ꩧ	ꩨ	ꩩ	ꩪ	ꩫ
ICC	54	55	56	57	58	59	60	61	62	63

Semi-vowels and Diphthong:

Latin EFEO		o	ai	i	i	im	ei	e	em	eng	u
Akhar Thrah	ꨩ	ꩺ	ꨠ	ꨪ	ꨫ	ꨬ	ꨭ	ꨮ	ꨯ	ꨰ	ꨱ
ICC	64	65	66	67	68	69	70	71	72	73	74
	â	r	i	u	l	lu	lu	ao	aong	aom	ng
	ꨲ	ꨳ	ꨴ	ꨵ	ꨶ	꨷	꨸	꨹	꨺	꨻	꨼
	75	76	77	78	79	80	81	82	83	84	85

2. Research Methodology

Research methodology plays an important role in the process of implementing the Project because it provides information and steps to carry out the program. In section 2 of literature reviews, we have presented a brief of historical research, systematic entire the documents related to language and Cham script such as ancient Cham script, traditional Cham script, modified Cham script and Cham Latin letters. As well as presented some Cham fonts in using and assigning Cham character. This section mainly explains the details of the research methods that are used to conduct this research. Parts of the study are presented in details as below

2.1 Research Design

This research implements quantitative approach. The variables will be measured by the tool so that the data can be analyzed using statistical procedures to generalize and systematic replication of results (Crewswell, 2008). When the research approach and research method was selected, we need to clearly define the target population for the study. Thereafter, selecting a sampling method ranging between probability and non-probability to determine the sample size was covered in the sample selection part (Trochim, 2000).

2.2 Population and Sample

A population imply to the whole group of individuals, events or objects having a common observable characteristic (M. Mugenda & G. Mugenda, 1999). The target population of this study included two groups of people group 1) university students, working people; and group 2) religious group of Cham in Vietnam. An estimated of Cham University students in Binh Thuan Province total about 120 people, and number of Cham working people in Binh Thuan are 195 people. The survey for religious group only conducted in three villages that are *Palei Canat, Palei Dik* and *Palei Aia Mamih*. The numbers of religious people in three villages were 105 people. Hence, the total of population in 2 groups was about 420.

In there, group University students, working people with a number of proportionate sample is 150 (75%) and group Religious people with a number of proportionate sample is 50 (25%). This is the total number of samples will be used to analyze and put into application.

Total number of items to be used for survey of Cham script is two items that are: (1) I choose Cham Akhar Thrah (traditional Cham script) to preserve, and (2) I choose Modified Cham Akhar Thrah to preserve. The overview is shown as in Table 2.

Table 2. Overview of Cham script survey

	Items	Questionnaire (face-to-face)	
Target group		Religious people	University students, working people
Cham script	1	✓	✓
	2	✓	✓

2.3 Cham Font Conversion Method

In order to convert Cham Latin EFEO to Akhar Thrah, our approach is to divide the process into two phases. 1) To convert from Cham Latin EFEO into intermediate characters code and each intermediate character code is corresponding only with a character of Cham Akhar Thrah. In this phase, the process is also selected from the list of special words. This solution can support converting in many different types of Cham Latin into many different Cham fonts. And 2) Convert from the intermediate character code corresponding to the Cham Akhar Thrah as the diagram below:
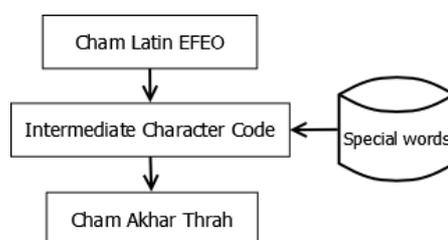

Figure 2. Diagram of converting Cham Latin EFEO to Cham Akhar Thrah script

2.3.1 Conversion EFEO Cham Latin to Intermediate Characters Code

In this phase, we divide the transition from Latin Cham to intermediate character code into 2 steps. Step 1: Convert from Latin Cham to intermediate character code based on Table 1, and Step 2: Continued performance the results from step 1 for these following cases.

i) Converting Takai Akhar

By convention transliteration of Cham Latin EFEO, the diacritics in this case always have position in front of vowels or syllables. Diacritics never have position in front of the character in end the word (*Akhar Matai*). Therefore, if the characters need to convert is not the first character of a word, and the convert string are not contain the *Langlikuk* word, the next character is not the final consonant (*Akhar Matai*), then proceed of conversion is according to the following table.

Intermediate Character	l(35)	r(34)	u(2)	i(1)
Convert to	l(79)	r(76)	u(78)	i(77)

ii) Converting Dar sa, Takai kuk

If the character need to be converted in the table below and their location is not the first character in a word, we convert as follows.

Intermediate Character	o(5)	u(2)
Convert to	o(65)	u(74)

iii) Converting é Vowel

If the intermediate character is *akhar é* (3), and the location is not the first character in a word, then convert to (o (65), e (71)).

iv) Convert Final Character (Akhar Matai)

If the intermediate character is the character in the table below and the location of character is end in a word, then convert as the table below:

Intermediate Character	k(6)	ng(10)	c(12)	t(19)	n(23)	p(26)	y(33)
Convert to	k(41)	ng(42)	c(43)	t(44)	n(45)	p(46)	y(47)
	r(34)	l(35)	w(36)	s(37)	h(38)	m(24)	
	r(48)	l(49)	w(50)	s(51)	h(52)	m(53)	

v) Convert Other Cases

If the intermediate character is the character i(1) and the location of character is end in a word then convert into i(69).

2.3.2 Convert from Intermediate Characters Code to Cham Script

To explore the preferable Cham script for preservation in section 2, we conducted a survey to choose which kind of Cham script need to preserve. With 200 samples for 2 groups: 1) Religious group and 2) University students and working group (source: Sang et al., 2013). The results show in Figure 3.
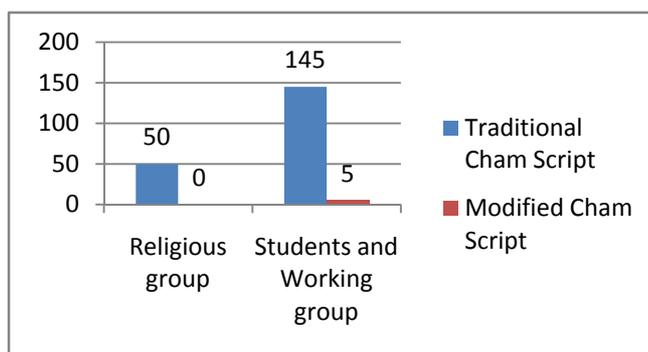

Figure 3. Choosing the preferable Cham script for preservation

The chart shows that the number of respondents in both two groups prefers to use the traditional Cham script (97.50%) whereas only 2.50% of the respondents supports for using the Modified Cham script. Besides, all the respondents from the religious group support the traditional Cham script; they do not know the modified Cham script. During the survey, they truly requested that the Cham people should use the Akhar Thrah (Traditional Cham script).

During the experiments for Cham font conversion, we have carry out the ADDIE method specifically in three steps which were develop, implement and evaluate. At each stage convert EFEO Cham Latin to intermediate character code with the cases such as converting *takai akhar*, converting *dar sa*, *takai kuk*, converting *é vowel*, convert final character (*akhar matai*), convert other cases, and especially the stage for checking and comparing in the Special words. We've been checking, evaluation and development of products good relatively.

In this experiment we have tested the program with three Cham works from EFEO Cham Latin to Cham Akhar Thrah (the traditional Cham script), and Cham script compare   in Cham-French dictionary of E. Aymonier & A. Cabaton; Cham-Vietnamese-French dictionary of G. Moussay and the results   is presented in Table 3.

Table 3. Result of Cham font conversion application

No.	Number of documents (Cham Poem)	Number of words	Checks (%)	Error (%)
1	Ariya Cam Bini	1823	100	0
2	Ariya Gleng Anak	2459	99.88	0.12
3	Nai Mai Mang Makah	2523	100	0

The following interface is the result of research products and tested the convert of three first sentence of the poem *Ariya Cam Bini* as presented in Figure 4.
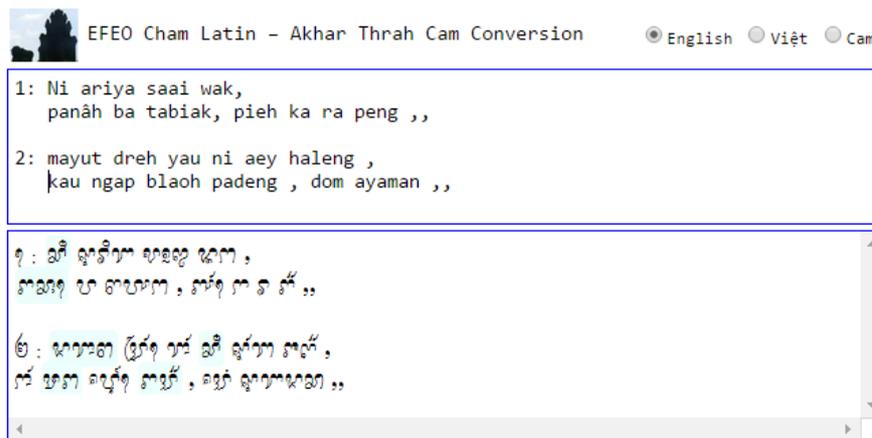

Figure 4. Interface of conversion EFEO Cham Latin to Cham Akhar Thrah

4. Results

Cham Font conversion has important implication in the studying, teaching and learning as well as preservation of Cham script and Cham language. To select Cham script conservation, we have examined as the results are presented in Figure 3. The chart shows that the number of respondents in all two groups supports using the traditional Cham script is 97,50% whereas is only 2,50% supports for using the Modified Cham script. Especially, there is 100% concern the Religious group supported the traditional Cham script.

The Cham font conversion method has significant important when convert the Cham Latin to Cham Akhar Thrah through intermedia character code and select the special words from the library.

Font conversion applications from EFEO Cham Latin to Cham Akhar Thrah with the interface as Figure 4. And we have checked the accuracy of three Cham Poems as presented in Table 3. Poem *Ariya Cam Bini* 100% (n=1823); *Ariya Gleng Anak* 99.88% (n=2459); *Nai Mai Mang Makah* 100% (n=2523). Cham font conversion application will be used in schools, institutions in the country and overseas as well as assist in teaching and learning Cham language.

5. Conclusions

We have presented a new approach of Cham font conversion in assiging of EFEO Cham Latin, Cham script character and the Intermediate Character Code (ICC). Font conversion approach includes two stages that are: 1). conversion of Cham Latin EFEO into intermediate characters code and 2) Conversion of intermediate characters code to Cham script. The experimental evaluation of Cham font conversion with different documents and assessment results are considered acceptable. Furthermore, the conversion is accurate on some common words. This experiment also has a small number of errors if the users enter words which are not from the Cham vocabulary or some proper nouns.

Font conversion applications bring significantly benefits for Cham community and in conversion of Cham language as well as applications in learning and teaching Cham language.

Acknowledgements

References

Crewswell, J. W. (2008). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative research* (3rd ed). Upper Saddle River: NJ:Merrill.

Grenoble, L. A., & Whaley, L. J. (2006). *Saving languages: An introduction to language revitalization*. Cambridge, UK: Cambridge University Press.

Mugenda, M., & Mugenda, G. (1999). *Research Methods. Quantitative and Qualitative Approaches*. Nairobi, Kenya.

Sang, V. N., Ali, M. B. B., & Said, H. B. (2013). *Information Technology and Indigenous language preservation, A case study in the Malayu Cham language in Viet Nam*. The 10-Year Implementation of the UNESCO Convention for the Safeguarding of Intangible Cultural Heritage: Experiences, Lessons and Future Directions in Hoi An City, Quang Nam Province.

Sastri, K. A. N. S. (1935). *L'origine de l'alphabet du Champa BEFEO, XXXV (1935)* (pp. 233-241). (Cham language and script in historical process). Conference Proceedings of the 2006 History of language and Cham script. 21-22 September. Kuala Lumpur, EFEO & Tokyo University of Foreign Studies, Kuala Lumpur, 2007 (CD-Rom).

Trochim, W. M. (2000). *The Research Method Knowledge Base* (2nd ed.). Cincinnati:Atomic Dog Publishing.

Copyrights