Full paper

Exploring the Cham Script to Develop the Transliteration Application for Cham Latin to Cham Script

Van Ngoc Sang[a]*, Mohamad Bin Bilal Ali[b], Noor Dayana Abd Halim[c]

aDepartment of Educational Technology, University Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

*Corresponding author: sangpodam@yahoo.com

Abstract

The main purpose of this research is to explore the preferable Cham script and Cham Latin; to develop application and evaluate the application products. Research sample is divided into two groups and the research instrument distributed into two forms of survey. The application was developed using ADDIE model. This survey, the acceptance of Cham script by religious group is 100%, Cham script online voting by student is 90.09%, Cham script by online questionnaire is 98.3%. EFEO Cham Latin by religious group is 100%, EFEO Cham Latin by online questionnaire is 95.4%. Viewed by sixteen experts using Fuzzy Delphi the results show that percentage of all items are 100%, more than what required (75%), the value of d for total construct is 0.02 (required ≤ 0.2). For conversion application, we have checked the accuracy percentage of four Cham poems and results *Ariya Gleng Anak* 99.88% (n=2459); *Nai Mai Mang Makah* 100% (n=2523); *Ariya Cam Bini* 100% (n=1823); Ariya Po Ceng 99.91% (n=2202). Using technologies to preserve the Cham language heritage is not only theoretically significant but also practically significant.

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

■1.0 INTRODUCTION AND BACKGROUND

Cham people are ethnic minorities group of Vietnam today, but Cham language is one of the languages in the family of multi-island "Austronesian". Therefore Cham language is close relationship with Malay language [1,2]. The Cham script is considered an endangered language in Vietnam. The crucial problem in Cham language is that the Cham Textbook Compiling Committee (CTCC) has not yet supported to preserve the traditional Cham script (TCS); moreover, they have made a new Cham script as the modified Cham script (MCS) for teaching and learning in primary school [3]. This is become emerge phenomenon and cause of the biggest conflict in Cham community. While most of Cham elders, students, intellectuals and Cham people really desire to conserve the TCS [4] because this not only the value heritage from the ancient but also the main key to open the door to understand the Cham Royal documents or old text materials which are written in TCS.

Transliteration application has been widely applied for some languages. The Hindi transliteration keyboard [5], the Arabic transliteration system [6] enables to type the Hindi or Arabic in English, instantly giving back the word written in Hindi or Arabic; The Japanese transliteration allows transliterate from English to Katakana, Hiragana ans Romaji [7]. For this study, we propose a Cham script transliteration method, which is based on the Cham Latin EFEO to Cham script (Cham *Akhar Thrah)*. Based on the above research on conflict and dangers of Cham script and language, we are interested in this issue and come out with three research questions of this study are as follows:

(i) What is the choice of Cham script among Cham people?
(ii) What is the choice of Cham Latin among Cham people?
(iii) What is the level of acceptance of the Cham font conversion application in term of use, convenience, and quality?

■2.0 METHODOLOGY

Research methodology play an important role in the process of implementing the project, it provides information and steps to carry out the program. In section 1, a briefly introduced of Cham language and script as well as unacceptable the modified Cham script teaching in primary school. Since then the conflict between the modified Cham script and traditional Cham script become one of the issues been discussed among educators. 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.

Population and Sample

Population of this study includes two groups of people group 1) university students, and government staff; and group 2) religious people of Cham in Vietnam. An estimated of Cham University students total number about 500 people, and Cham government staff are 500 people. The scope of the survey for religious group only performed in three villages that are Palei Aia Mamih, Palei Dik and Palei Panat. For each village the numbers of religious people are 37, 22, and 21 respectively. Hence, the total of population in two groups counted about 1080. To draw conclusions about the population from the sample, the sample size of any research must be determined in the phase design research. However, before determining the size of the sample needs to be drawn from the population, several factors must be considered. According to [8], the sample size representative of two groups is 284. The details are shown in Table 1.

Table 1. Proportionate sample techniques of group study

Group of study	Population	Proportionate sample	(%)
University students and government staff	1000	263	93%
Religious people	80	21	7%
Total	1080	284	100%

Based on the Table 1 the required sample size for this study has been stated at 284. In which 93% (263) were come from group 1 while 7% (21) were comes from group 2.

Transliteration Application Analysis

Font conversion application has been widely applied by many researchers to convert different font formats for different purposes. Here, we propose two Cham font conversion methods, which is applied for Cham font conversion from Cham Latin EFEO to Cham script as shown in Figure 1.

Algorithm 1: Advantage: The Cham font converted exactly by transformation rules. Disadvantage: Incorrect conversion of some special words (irregular).

Algorithm 2: Advantage: The Cham font converted exactly by transformation rules. Disadvantage: Fix the bug in the algorithm one; the special word (irregular) is converting well. However, in order for the program to be applied well, the need to add more special words into the list of special words.
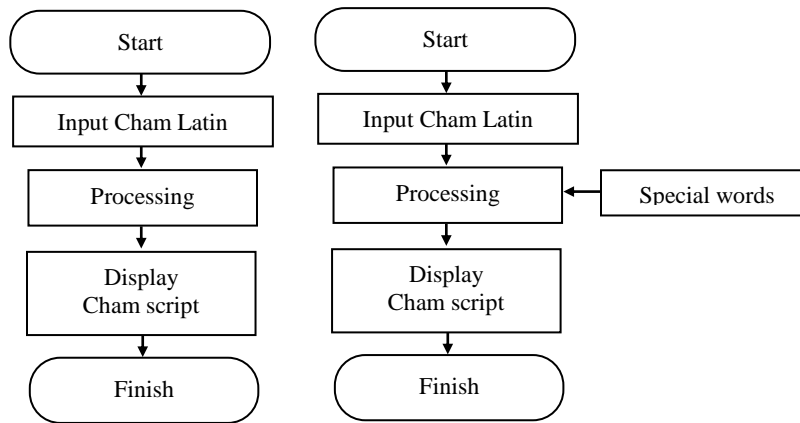

Figure 1. Algorithm 1 and Algorithm 2

Cham Font Transliteration Method

In order to convert EFEO Cham Latin to Cham Akhar Thrah, our approach is to divide the process into two phases [9]. First to convert from EFEO Cham Latin 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 selecting the word from Special words list and Special meaning list. This solution can support conversion of many different types of Cham Latin into various Cham fonts. Meanwhile the second phase is converting the intermediate character code corresponding to the Cham Akhar Thrah as shown in the Figure 2.
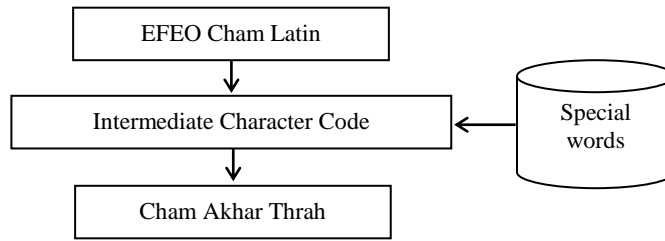
Figure 2. Diagram of converting EFEO Cham Latin to Cham Akhar Thrah script

Cham Font Transliteration Design

In order to design the interfaces for Cham font conversion application, we develop the general design consists of a textbox to input the source text to be converted, and a textbox to display the conversion results. In addition, we also design the OK, Copy and the Size button. See Figure 3.

Figure 3. Interface of conversion and select word has many options

■3.0 EXPERIMENTS

In this part, we have survey the Cham script and Cham Latin through two groups of people group 1) university students, and government staff; and group 2) religious people. Simultaneous to evaluate the Acceptant Level of application products from respondents and experts.

Exploring the Preferable Cham Script

This part will address the research question one: What is the choice of Cham script among Cham people? To answer this question we had survey in three different ways which are survey face-to-face to religious groups, online voting survey and online questionnaire by Google Docs to university students and government staff group.

Firstly, finding the preferable Cham script based on the data collections face-to-face from religious people group, the results showed that most of the religious people group choosing the traditional Cham script to preserve. The results are display as in Table 2.

Table 2. Selecting the Cham script by religious group

Item	Total	Percentage
Traditional Cham script	80	100%
Modified Cham script	0	0%

respondent, specifically, the religious people group has confirmed to choose the traditional Cham script. Besides, the modified Cham script with no choice. This implies that the entire religious group has lowest perception towards the modified Cham script.

Secondly, findings the preferable Cham script based on online voting. The survey was conducted on website gulpataom.com. The results showed that most of the university student and government staff choosing the traditional Cham script to preserve. The results are presented as in Table 3.

Table 3. Selecting the Cham script by online voting

Item	Total	Percentage
Traditional Cham script	600	90.09%
Modified Cham script	47	7.06%
Missing	19	2.85%
Total respondent	666	100.00%

Based on Table 3, the total result showed that the number of respondent selected the traditional Cham scripts to preserve is 600 (90.09%). This means that the number of those surveyed were aware of the traditional Cham script need to continue to maintain, preserve and develop. Whereas, the total number of respondents selected the modified Cham script are only 47 (7.07%). This means that the Cham community has disagreed about the modified Cham script.

Thirdly, findings the preferable Cham script based on online questionnaire by Google Docs was conducted by university student and government staff groups. The respondent selecting the traditional Cham script or modified Cham script to preserve by answers agree or disagreed. Results are presented in Table 4.

Table 4. Selecting the Cham script by online questionnaire

Item	Agreed	Disagreed	Total
Traditional Cham script	178 (98.3%)	3 (1.7%)	181 (100%)
Modified Cham script	17 (9.8%)	157 (90.2%)	174 (100%)

Based on the table 4, the results showed that the number of people agreed to choose the traditional Cham script to preserve has achieved 178 (98.3%), and the number disagreed has only 3 (1.7%). In contrast, the number of people agreed to choose modified Cham script is only 17 (9.8%), and the number of people disagree has achieved 157 (90.2%). These results tell us that the respondents of group university students and government staff have a good perception towards selected the traditional Cham scripts to preserve. Meanwhile, the modified Cham script agrees with a lower value. This implies that among the respondent of university students and government staff group has the lowest perception towards the way to select the modified Cham script.

Through three surveys about selecting the Cham script to preserve, results showed that the religious group has selected the traditional Cham script to preserve is 80 (100%). University students and government staff via online voting survey has selected the traditional Cham script is 587 (90.87%) and via online questionnaire survey by Google Docs is 173 (98.3%). These results tell us that all the Cham religious group, university students, intellectuals and Cham community has a good perception towards selected the traditional Cham script to preserve. Conversely, the modified Cham script was selected very low. This implied that among the Cham community, the respondent has the lowest perception towards the way to select the modified Cham script to preserve.

Exploring the Preferable Cham Latin

This part will address the research question two: What is the choice of Cham Latin among Cham people? To answer this question we had survey in two different ways which are survey face-to-face to religious groups, and online questionnaire by Google Docs to university students and government staff group.

Firstly, finding the preferable Cham Latin based on the data collections face-to-face from religious people group, the results showed that most of the religious people group choosing the EFEO Cham Latin (the Latin based on traditional Cham script) to preserve. The results are presented as in Table 5.

Table 5. Selecting the Cham Latin by religious group

Item	Total	Percentage
EFEO Cham Latin	80	100%
CTCC Cham Latin	0	0%

Based on the Table 5. The results showed that the total of religious group selecting the EFEO Cham Latin is 80 (100%). This implied that the religious group has a good perception towards preserve the traditional Cham script. Hence, they also select the EFEO Cham Latin which based on the traditional Cham script to preserve.

Secondly, findings the preferable Cham Latin based via questionnaire online by Google Docs were conducted by university student and government staff groups. The respondent selected the answers agree or disagreed for selecting the EFEO Cham Latin or CTCC Cham Latin to preserve. Results are presented in Table 6.

Table 6. Selecting the Cham Latin by online questionnaire

Item	Agreed	Disagreed	Total
EFEO Cham Latin (Based on traditional Cham script)	166 (95.4%)	8 (4.6%)	174 (100%)
CTCC Cham Latin (Based on modified Cham script)	11 (6.6%)	155 (93.4%)	166 (100%)

Based on the Table 6, the results showed that the number of people agreed to choose the EFEO Cham Latin to use has achieved 166 (95.4%), and the number of people disagreed has only 8 (4.6%). In contrast, the number of people agreed to choose the CTCC Cham Latin is only 11 (6.6%), and the number of people disagree has achieved 155 (93.4%). These results tell us that the respondents of group university students and government staff have a good perception towards selected the EFEO Cham Latin to preserve. Meanwhile, the CTCC Cham Latin agrees with a lower value. This implies that among the respondent of university students and government staff group has the lowest perception towards the way to select the CTCC Cham Latin.

Through two surveys about selecting Cham Latin, results showed that the religious group has selected the EFEO Cham Latin is 80 (100%). University student and government staff via online questionnaire survey by Google Docs with the highest value is 162 (95.3%). This is proves that all the Cham religious group, university students, and government staff have a good perception towards selected the EFEO Cham Latin to use. Conversely, the CTCC Cham Latin was selected very low is only 11 (6.8%). This implied that among the Cham community, the respondent has the lowest perception towards the way to select the CTCC Cham Latin.

Acceptance Level of Transliteration Application

This section will address the level acceptance of the Cham font conversion application in term of use, convenience, and quality and be evaluated by two groups are Respondents and Experts

Acceptance Level from Respondents

This first part will address the level acceptance of the Cham font conversion from respondents. Each question in this part with a scale from 1 to 5 is used to measure the evaluation of Cham font conversion through their answer in the questionnaire. The assessment result for each question is presented in Table 7.

Table 7. The level acceptance of Cham font conversion

No	Item	Agree	Strongly Agree	Total
1	This application convert exactly from EFEO Cham Latin to traditional Cham script	10 (18.52%)	44 (81.48%)	54 (100%)
2	I like the way to select one word has many options for Cham script	17 (31.48%)	37 (68.52%)	54 (100%)
3	I like the way to select the word has many meanings for Cham script	21 (38.89%)	33 (61.11%)	54 (100%)
4	Same format when copy from application paste to MS Word	14 (25.93%)	40 (74.07%)	54 (100%)

Based on the Table 7, for all questions the results showed that the respondent answer with two choices are strongly agreed and agreed with the total is 54 (100%). This implied that the respondent has a good level of acceptance towards Cham font conversion. All the respondent only stated their agreeability on agree and strongly agree scale and no one stated their disagreeableness. Specifically, from the four items asked, the item number 1 showed the highest value of respondent choosing the answer strongly agree is 44 (81.48%). This result tells us that the respondent has a good perception towards this application convert exactly from EFEO Cham Latin to traditional Cham script. Besides, the item with lower value is item number 3. This implies that among the four item, respondent has the lowest perception towards the way to select the word has many meanings for Cham script. However, this item still consider as high value. To analyses we compute the score of four items for each respondent. The result is displayed as in Table 8.

Table 8. Respondent and total score of Cham font conversion

Number of respondent	Total score
24	20
13	19
8	18
3	17
6	16

In order to make decision, we decided to categorize the score to two categories as follow:

Table 9. Categories of total score for Cham font conversion

Categories	Total score
Not Accepted	4-12
Accepted	13-20

Based on Table 8 and Table 9, it can be concluded that all respondent 54 (100%) accept the Cham font conversion.

Acceptance Level from Experts

This second part will address the acceptance level of the Cham font conversion from experts. We selected a total of 16 experts to evaluate four items instrument using Fuzzy Delphi. For each item in this instrument with a scale from 1 to 5. The level acceptance from Experts is presented in Table 10.

Table 10. Threshold Value (d) and Percentage Consensus by Experts of Cham Font Conversion

Experts	Items			
	Item 1	Item 2	Item 3	Item 4
1	0.01	0.01	0.01	0.00
2	0.01	0.01	0.01	0.00
3	0.01	0.01	0.01	0.00
4	0.01	0.01	0.01	0.00
5	0.01	0.01	0.01	0.00
6	0.19	0.01	0.01	0.00
7	0.01	0.01	0.01	0.00
8	0.01	0.01	0.01	0.00
9	0.01	0.01	0.01	0.00
10	0.01	0.01	0.01	0.00
11	0.01	0.19	0.19	0.00
12	0.01	0.01	0.01	0.00
13	0.01	0.01	0.01	0.00
14	0.01	0.01	0.01	0.00
15	0.01	0.01	0.01	0.00
16	0.01	0.01	0.01	0.00
Frequency $d \leq 2$	16.00	16.00	16.00	16.00
Percentage item $d \leq 2$	100%	100%	100%	100%
Value d for total construct	0.02			

From Table 10, the threshold ($d_{m,n}$) for each item based on the expertise and the percentage expert consensus ($d \leq 0.2$) for all four items are 100%, more than what required (75%). The value of d for total construct is 0.02 (required $d \leq 0.2$). This can be concluded that all 16 experts has come to consensuses, the Cham font conversion is acceptable.

■4.0 RESULTS

Based on the survey results on selecting the Cham script to preserve, we built the conversion application for Cham Latin to Cham script. In the experiment the program has been tested with four Cham Poems from EFEO Cham Latin to Cham script (the traditional Cham script), and Cham script has compared with two resources by Cham - French dictionary of E. Aymonier & A. Cabaton; Cham - Vietnamese - French dictionary of G. Moussay and the results of the accuracy percentage of four Cham Poems is presented in Table 11.

Table 11. Result of Cham font conversion application

Number of documents (Cham Poem)	Number of words	Checks (%)	Error (%)
Ariya Gleng anak	2459	99.88	0.12
Nai mai mang Makah	2523	99.92	0.08
Ariya Cam Bini	1823	100.00	0.00
Ariya Po Ceng	2202	99.91	0.09

Based on the Table 11, the Ariya Gleng anak 99.88% (n=2459); Nai mai mang Makah 99.92% (n=2523); Ariya Cam Bini 100% (n=1823); Ariya Po Ceng 99.91% (n=2202). Cham font conversion application will be used in schools, institutions as well as in assisting teaching and learning Cham language.

■**5.0 CONCLUSIONS**

In this paper, we have presented a new approach. First we have been conducted the survey to identify the choice of Cham script and Cham Latin among Cham people. Second, we proposed a technique Cham Font Conversion model to convert the EFEO Cham Latin to Cham script. And third, 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 which are not from the Cham vocabulary or some proper nouns. Further research will develop applications both online and offline and ensure that Cham language is highly valued and continues to be developed in future.

References

[1]   Coedes, G. (1939). La plus ancienne inscription en langue chame. Eastern and Indian Studies in Honour of F.W. Thomas. New Indian Antiquary Extra Series I. Po Dharma, " History of language and Cham script",Conference *Proceedings on 21-22 September, 2006*. Kuala Lumpur, EFEO & Tokyo University of Foreign Studies, Kuala Lumpur, 2007 (CD-Rom). p. 46-49.
[2]   Lafont, P. B. (2011). Vuong Quoc Champa: Dia Du, Dan Cu va Lich Su (Kingdom of Champa: Geography, Population and History) (Vol. Champaka No. 11). *San Jose, USA: International Office of Champa*.
[3]   Dharma, P. (2006). Ngon ngu chu viet Cham trong qua trinh lich su (Cham language and script in historical process). In 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).
[4]   Phan, T. (2006). Ngon ngu chu viet Cham trong qua trinh lich su (Cham language and script in historical process). *In 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).
[5]   KeyNounce**,** (2014). *Hindi Transliteration Keyboard*. Retrieved on Oct 20, 2016, from https://itunes.apple.com/in/app/hindi-transliteration-keyboard/id968974271?mt=8.
[6]   LLC**,** (2010). Arabic transliteration. Retrieved on Oct 20, 2016, from http://www.arabic-keyboard.org/arabic/arabic-transliteration.php.
[7]   Orthogonal, LLC. (2010). *Japanese transliteration*. Retrieved on Oct 20, 2016, from http://japanesetransliteration.com/
[8]   M Krejcie, R. V. and D. W. Morgan (1970). Determining Sample Size for Research Activities. *Educational and Psychological Measurement* (30) 607-610.
[9]   Sang Ngoc,V. and Mohamad Bin Bilal Ali (2014). Technology Application in Cham Language Conservation. *The 1st International Education Postgraduate Seminar (IEPS2014)*. 22-23, November. Johor, Malaysia.