

TECHNOLOGY APPLICATIONS IN CHAM LANGUAGE CONSERVATION

Van Ngoc Sang, Mohamad Bin Bilal Ali
Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia
sangpodam@yahoo.com, mba@utm.my

Abstract: All the ethnic language is established and developed for a long time in the historical process. Cham language is the indigenous language in Vietnam but now it becomes the minority language. However, the Cham script was appeared in 4th century on stone stele in Tra Kieu. It was considered as the first language in South East Asia. Due to many reasons and influenced by history, Cham population are now just over 145 thousand people, therefore the Cham language is facing endangered and they are at risk of deformation for various reasons.

This paper aims to introduce some specific results of information technology in Cham language heritage conservation in Vietnam such as creating Cham fonts, and Cham font conversion. In this experiment, for Cham font, the acceptable level of Cham Thrah font is 60% of usability and 65% of high accuracy. On the other hand, the acceptable level for *champa* font only is (12%, 8%), *camtanran* is (5%, 14%) and *bingu tanran* (23%, 13%). For Cham font 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). Cham font and Cham font conversion are necessities and meaningful in conservation of Cham script. Using technologies to preserve the Cham language heritage is not only theoretically significant but also has practical significance conservation of heritage languages of ethnic groups in the development process.

Keywords: cham language; cham script; cham font; cham conversion; cham font conversion.

1. Introduction and Background

Cham people are ethnic group of the Champa kingdom was established in the second century in central Vietnam. Cham language is one of the languages in the family of multi-island "Austronesian". Therefore Cham language is closely related with Malay language (Dharma, 2006). As compared to the Malay language, researchers have found traces on stone stele in Java island in the 7th century (Casparis, 1956), while Cham letters has also appeared on the first stone stele, namely of Dong Yen Chau (Tra Kieu) in the 4th century (Coedes, 1939; Lafont, 2011).

Using information technology (IT) to preserve the endangered language is significant in the current, when the digital technology has developed with the diverse features of the storage and dissemination of information. IT has contributed in the indigenous language education, restoring and preserving of ancient script system is a very obvious fact in many countries (Grenoble & Whaley, 2006).

Using font to type Cham script on the computer has become essential issues to help compiling textbooks, data storage, teaching and learning Cham script. Cham-Tanran font was created by the EFEO for Macintosh keyboard in 1988 showed that this font has typography and typeface relatively standard. However, this font is still having some defective and technical errors. *Cam Thrah.ttf* font created by Sang (2002, 2012) and *Bingu di tanran.ttf* by Cham Unesco (2012) developed from Cham-Tanran (EFEO) rebuilt for Windows keyboard. However these Cham fonts inconsistent the value of the code page, this causes many difficulties in the exchange of information. Besides, these fonts are still limited in the technical thereby create documents not ensure scientific and aesthetic.

Based on the above research on conflict and dangers of the Cham language, as well as the use inconsistent of Cham font typing and storing, we are interested in this issue and the three research questions of this study are as follows:

- (i) What is the choice of Cham script among Cham people?
- (ii) How to develop application technology for Cham's script preservation (such as Cham fonts, and convert Cham font)?
- (iii) What is the acceptable level of the Cham font conversion application?

2. Research 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 brief of historical research on

Cham script and Cham fonts are presented and used over time to preserve and conserve the Cham Akhar Thrah.

2.1 Population and Sample

The target population of this study included two groups of people 1) university students, working people; and 2) religious people. An estimated of Cham University students in Binh Thuan Province total about 120 students, and number of Cham working people in Binh Thuan are 195. The scope of the survey for religious group only performed 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 is 420. Meanwhile the totals of items used for Cham script survey are two items: Choosing traditional Cham scrip; and Choosing modified Cham scrip.

According to (Krejcie & Morgan, 1970), the ever increasing demand for research has created a need for an efficient method of determining the sample size needed to be presentative of a given population. In order to know the number of sample size to be studied in a total 420 people of population in two groups. To obtain the required sample size using the formula at population is 420, the sample size representative of two groups is 200. The details are shown in Table 1.

Table 1 Proportionate sample techniques of group study

Group of study	Population	Proportionate sample	(%)
University students	315	150	75%
Religious people	105	50	25%
Total	420	200	100%

Based on the Table 1 the required sample size for this study has been stated at 200.

Out of 200, 75% (150) were come from group1 while 25% (50) were comes from group 2.

2.2 Cham Font Design

In order to design new *Cham* fonts, we based on a number of valuable documents and dictionary which are using today include Cham-French dictionary of E. Aymonier & A. Cabaton; Cham-Vietnamese-French dictionary of G. Moussay, and *Akhar Thrah Cham* materials used in Cambodia that is currently stored in the library in French.

Akhar Thrah Script is a script derived from the Sanskrit, difference Latin characters. To design the font for *Akhar Thrah*, the first step is analysing the number of characters needed to create the font for *Akhar Thrah* script. The total numbers of *Akhar Thrah* characters required design are 86 characters. After analyzing and removing some of the same characters as (H), (3), (o), (r), respectively (H) 1, (i) 3, (o) 5, (6) 6. Thus, the total number of characters needed to be designed for *Akhar Thrah* fonts are 82 characters. However, for the convenience of typing Cham script, we design and add seven new punctuation and one characters for using in Latin system, which is 𑜀𑜂𑜆𑜐, 𑜀𑜂𑜆𑜑, 𑜀𑜂𑜆𑜒, 𑜀𑜂𑜆𑜓, 𑜀𑜂𑜆𑜔, 𑜀𑜂𑜆𑜕, 𑜀𑜂𑜆𑜖. Then, the total numbers of Cham script character need to design are 90 characters as show in Figure 1.

The screenshot shows a software window titled "Cham Thrah1". It has a menu bar with "View By: Unicode", "Hex: 20", and "Keystroke:". There is a search bar labeled "Search by text ranges". Below the menu bar is a grid of characters. The grid has 16 columns and 10 rows. The first column contains code points from 0020 to 0070. The subsequent columns contain the corresponding Cham script characters. The characters are arranged in a grid that matches the data in Figure 1.

0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
	𑜀𑜂𑜆𑜐	𑜀𑜂𑜆𑜑	𑜀𑜂𑜆𑜒	𑜀𑜂𑜆𑜓	𑜀𑜂𑜆𑜔	𑜀𑜂𑜆𑜕	𑜀𑜂𑜆𑜖	𑜀𑜂𑜆𑜗	𑜀𑜂𑜆𑜘	𑜀𑜂𑜆𑜙	𑜀𑜂𑜆𑜚	𑜀𑜂𑜆𑜛	𑜀𑜂𑜆𑜜	𑜀𑜂𑜆𑜝	𑜀𑜂𑜆𑜞
0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
𑜀𑜂𑜆𑜟	𑜀𑜂𑜆𑜠	𑜀𑜂𑜆𑜡	𑜀𑜂𑜆𑜢	𑜀𑜂𑜆𑜣	𑜀𑜂𑜆𑜤	𑜀𑜂𑜆𑜥	𑜀𑜂𑜆𑜦	𑜀𑜂𑜆𑜧	𑜀𑜂𑜆𑜨	𑜀𑜂𑜆𑜩	𑜀𑜂𑜆𑜪	𑜀𑜂𑜆𑜫	𑜀𑜂𑜆𑜬	𑜀𑜂𑜆𑜭	𑜀𑜂𑜆𑜮
0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
𑜀𑜂𑜆𑜯	𑜀𑜂𑜆𑜰	𑜀𑜂𑜆𑜱	𑜀𑜂𑜆𑜲	𑜀𑜂𑜆𑜳	𑜀𑜂𑜆𑜴	𑜀𑜂𑜆𑜵	𑜀𑜂𑜆𑜶	𑜀𑜂𑜆𑜷	𑜀𑜂𑜆𑜸	𑜀𑜂𑜆𑜹	𑜀𑜂𑜆𑜺	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾
0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	005A	005B	005C	005D	005E	005F
𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼
0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	006A	006B	006C	006D	006E	006F
𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼
0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F
𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼	𑜀𑜂𑜆𑜽	𑜀𑜂𑜆𑜾	𑜀𑜂𑜆𑜿	𑜀𑜂𑜆𑜻	𑜀𑜂𑜆𑜼

Figure 1. Code point location of Cham Thrah font

2.3 Cham Font Conversion Method

In order to convert EFEO Cham Latin to Cham Akhar Thrah, our approach is to divide the process into two phases. 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.

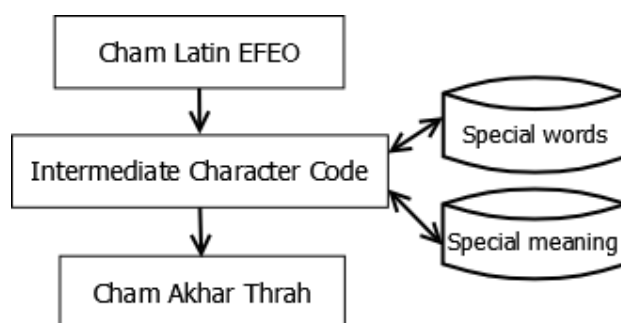


Figure 2. Diagram of converting EFEO Cham Latin to Cham Akhar Thrah script

3. EXPERIMENTS

3.1. Exploring the Preferable Cham Script

To explore the preferable Cham script for preservation, a survey to choose which kind of Cham script to use for Cham community. With 200 samples for 2 groups: 1) Religious group and 2) University students and working group (Sang, Ali, & Said, 2013). The results show in Figure 3.

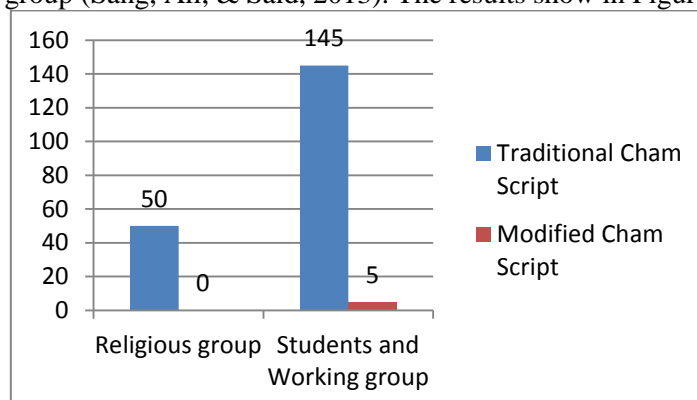


Figure 3. Chart of choosing Cham script to use

The chart shows that the number of respondents in all two groups supports using the traditional Cham script is 97,50% whereas only 2,50% supports for using the Modified Cham script. Suprisingly, there is 100% of the Religious group supported the traditional Cham script. During survey, they requested that the Cham people should use the traditional Cham script.

3.2 Creating Cham Font

Fontographer was use to create favorable conditions for design Cham font with a number of important functions such as viewing all the character designs in the window, test or modify the font encoding and kerning in Metrics Window,...and supported on many different operating systems such as Macintosh, Linus, or Windows. This font has been built based on Cham dictionary and design on Unicode. *Cham Thrah* font typeface is based on drawings presented in Cham - Vietnamese - French dictionary of G. Moussay, published in 1971. The result is shown in Figure 1.

3.3 Converting Cham Font Evaluation

In this experiment the program has been tested with four Cham Poems from EFEO Cham Latin to Cham Akhar Thrah (the traditional Cham script), and Cham script has compared by Cham - French

dictionary of E. Aymonier & A. Cabaton; Cham - Vietnamese - French dictionary of G. Moussay and the results is presented in Table 2.

Table 2. 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	0
Ariya Po Ceng	2202	99,91	0,09

The following interface is the result of research products and tested the conversion of two first sentence of the poem Nai mai mang Makah as presented in Figure 4.

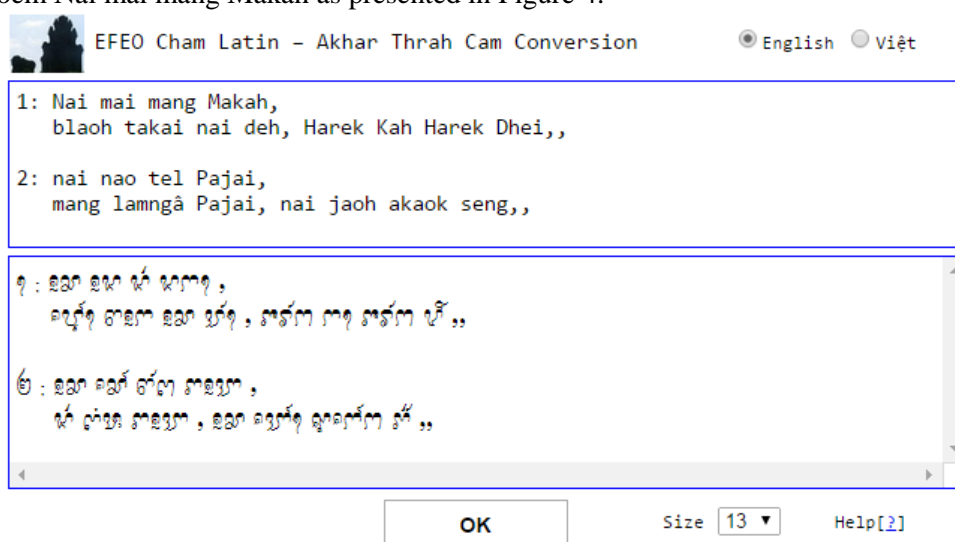


Figure 4. EFEO Cham Latin to Cham Akhar Thrah interface conversion

4. RESULTS

In the aspect of Cham script conservation, 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 only 2,50% supports for using the Modified Cham script. Besides, there is 100% concern the Religious group supported the Traditional Cham script.

Cham Thrah fonts has significant in using, typing document, research, communication and teaching Cham language. Compared with *camtanran.ttf*, *champa.ttf*, *bingudi tanran.ttf*, it is so that *cham Thrah* fonts have outperformed with high accuracy and easy to use. Moreover, on both Windows and Macintosh keyboard, the cham thrah font has the same codepoint value, and then it is very convenient for users. In survey 50 users, result shows that *cham thrah* font has surpassed in accuracy and ease of use as shown in Figure 5. Specifically, the acceptable level is 60% of usability and 65% of high accuracy. On the other hand, the acceptable level for *champa* font only is (12%, 8%), *camtanran* is (5%, 14%) and *bingu tanran* (23%, 13%).

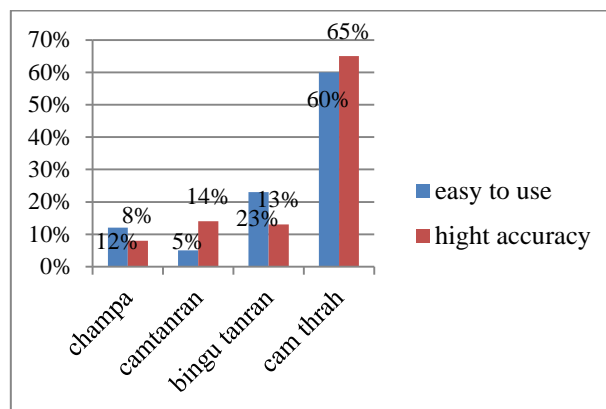


Figure 5. Chart of survey and assessment Cham font

Cham font conversion applications with the interface are shown in the Figure 4. The accuracy percentage of four Cham Poems is presented in Table 2. Poem Ariya Gleng anak (99,88%); Nai mai mang Makah (99,92%); Ariya Cam Bini (100%); Ariya Po Ceng (99,91%). Cham font conversion application will be used in schools, institutions as well as in assisting teaching and learning Cham language.

5. CONCLUSIONS

This study has been conducted to identify the choice of Cham script among Cham people about traditional Cham script and modified Cham script, and also carry out a new approach in analysing and designing all code point value of Cham fonts reasonable on the keyboard. Cham fonts have been design based on Unicode and using for both Macintosh and Windows keyboard. Hence, Cham Thrah character and Times New Roman characters have the same size in the same text. Therefore the conversion or storage information on both operating systems is stable results. In addition, the applications of Cham font conversion from EFEO Cham Latin to Cham Akhar Thrah from the experiments were highly accurate. Information technology applications to create Cham font and converting Cham font have an important implication in Cham language conservation.

REFERENCES

- Casparis, J. G. d. (1956). *Prasasti Indonesia II: Selected Inscriptions from the 7th to the 9th Century A.D.* Bandung: A.C. Nix & Co.
- 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.
- Dharma, P. (2006). *Ngon ngu chu viet Cham trong qua trinh lich su (Cham language and script in historical process)*. Paper presented at the 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).
- Grenoble, L. A., & Whaley, L. J. (2006). Saving languages: An introduction to language revitalization. *Cambridge, UK: Cambridge University Press.*
- Krejcie, R. V., & Morgan, D. W. (1970). Determining Sample Size for Research Activities. *Educational and Psychological Measurement, (30)* 607-610.
- 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.
- Sang, V. N., Ali, M. B. B., & Said, H. B. (2013). Information Technology and Indigenous language preservation, A case study in the Malayyu Cham language in Viet Nam. *International conference "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.*