International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I10P226

Volume 69 Issue 10, 204-210, October, 2021
©2021 Seventh Sense Research Group®

Digitizing the Minority Language Documents in Vietnam by Using Unicode

Hoang Thi My Le

Hoang Thi My Le & University of Technology and Education-The University of Danang, Vietnam, 48, Caothang, Danang, Vietnam

[1] htmle@ute.udn.vn

Abstract - *Using its own fonts in documents of Vietnamese ethnic minority languages is a major obstacle for digitization to develop information systems. Therefore, Vietnamese ethnic minority language documents face difficulties in displaying, storing, processing, and exchanging on the internet or between computers that do not have the same font. These difficulties have affected the digitization to develop the information system of ethnic minority areas in Vietnam. In order to overcome the above difficulties, the paper proposes a solution for encoding the Unicode character sets of ethnic minority languages in Vietnam. This solution is applied in language processing of the Ede ethnic minority in Vietnam, specifically: using Unicode font in documents and converting documents using own fonts to Unicode fonts*

Keywords — *Unicode, Encoding, natural language processing, minority language processing, Unicode font.*

I. INTRODUCTION

All information processing activities in computers are related to a text editor, such as communication between people and computers in applications, utilities or web services, programming, etc. There are many different purposes in a text editor, such as social communication, administrative activities in the office, storing and searching documents, building information systems, translating natural languages, applications in artificial intelligence, printing, electronic publishing, etc.

The basic operations in a text editor are typing characters, formatting text, and selecting fonts. There are also other common operations as creating a new document in the format, opening existing documents to editing, converting, sharing, or exchanging that often face difficulties with documents without Unicode fonts [1].

In the text editor, using different character codes in the country is the major obstacle in the information system development [2], [3]. If users have a habit of using their own font in the text, then it will not affect displaying text on their computer. However, using own fonts in documents will be a major obstacle for exchanging information at the national and international level. The documents will not be readable with strange characters because the corresponding font is not installed on the computer.

Unicode has overcome incompatibilities between the character codes and created the standard character code for all languages in the world [4], [5], in which there are Vietnamese, Korean, Japanese, Thai, Chinese characters, etc. It does solve not only the displaying documents but also paves the way for developing the problems in natural language processing such as spelling and grammar checking, etc. [6], [7].

II. DISPLAYING THE CHARACTER SETS OF ETHNIC MINORITY LANGUAGES IN VIETNAM

Currently, Vietnam has 29 ethnic minority groups with written characters and 26 without written characters. There are 21 ethnic minorities with Latin scripts and 8 with ancient scripts [7], [8]. Researching on displaying Vietnamese character sets has been implemented quite early and had successful results. In addition, displaying the character sets Of Ethnic Minority Languages also had the following results:

Displaying Cham language, the researcher used ASCII for encoding, the Corel Draw software to draw 65 Cham language characters, and the FontCreator software to create its own fonts. Cham language text editor uses created Cham fonts and Vietnamese typing tool [9], [10].

The Ede language font set is created in the Fontographer software by overwriting some characters in the Unicode fonts [11].

Displaying Ede, GiaRai, BaNa, XoDang, and M'Nong languages, the researchers used ASCII for encoding accented characters and created TayNguyenKey font set by the Fontographer tool. Editing text of Ede, GiaRai, BaNa, XoDang, and M'Nong languages uses the created font set and Vietnamese typing tool.

The VnKey typing tool has its own font set and supports to display Ede, GiaRai, M'Nong, CoHo, XoDang, SanChay languages, and Vietnamese [13].

Displaying the Thai Son La language, the researchers built their own fonts by replacing the keyboard characters with Thai language characters. The own fonts does not depend on typing tool because the number of Thai characters is less than the number of keyboard characters [14].

The research results on displaying the character sets of the ethnic minority languages have the advantages and disadvantages:

Advantages

- Editing the documents of the ethnic minority languages;
- Solving the problem of displaying Cham, Thai, Ede, GiaRai, BaNa, M'Nong, CoHo, XoDang, and SanChay languages;
- Spreading IT applications and new technological scientific to ethnic minorities.

Disadvantages

- The research results are not systematic, not clearly oriented, still disconnected, not sharing and serving only minority communities in each district.
- Unicode fonts have not been used in the text editor.

Currently, editing minority language documents in applications need to be interested researching to digitizing for the developing information system of minority communities in general and Vietnam minority communities in particular.

III. THE SOLUTION USING UNICODE FONT IN MINORITY LANGUAGE DOCUMENTS

Using Unicode font in the Vietnamese minority language (VML) text editor is proposed basing on the most of 21 Vietnam ethnic minorities have Latin character sets [15], with accents and vocals similar to Vietnamese. Some accented characters are different, but not much. The following criteria are set up in the solution:

- Applying for VML with Latin character set,
- Using Unicode to encode character sets,
- Inheriting Vietnamese typing tool.
- The workflow diagram of the solution is shown in Fig. 1.
- Step 1: Splitting VML character set in following three groups:
  - Group 1: The characters are in the Vietnamese character set.
  - Group 2: The characters are not in the Vietnamese character set but not in Unicode.
  - Group 3: The characters are neither in Unicode nor in the Vietnamese character set.
- Step 2: Mapping characters of groups 2 and 3 to Unicode for determining the corresponding hexadecimal values.

  Image of x through mapping f, y=f(x), x ∈{X} and y∈{Y};

  In which X is the character set in groups 2 and 3; Y is the set of hexadecimal values in Unicode. Example: H014F= f(ŏ)
- Step 3: Stipulating for typing method of characters in.
- Step 4: Creating an environment to install hexadecimal values and stipulated typing methods on the WinVNKey application [16].

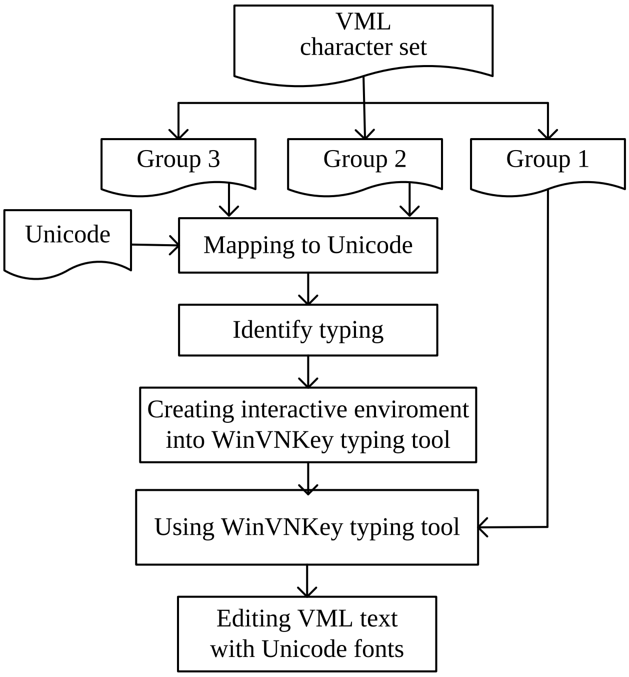

VML character set
Group 3
Group 2
Group 1
Unicode
Mapping to Unicode
Identify typing
Creating interactive enviroment into WinVNKey typing tool
Using WinVNKey typing tool
Editing VML text with Unicode fonts

Fig. 1: The workflow of using Unicode to encode VML character set

This research result has improved the restrictions of the previous research result that was presented at The Fourth International Conference on Knowledge and System Engineering [17]. These improvements are shown in Table 1.

TABLE I: COMPARING THE PAPER RESULT WITH THE PREVIOUS RESULT

Improvement	Result of the paper	Previous result
Characters in group 3	8	10
Applying	All text editor application	WinWord application

This solution has contributed to solving the problem of using Unicode font in VML text editor applications. This is the infrastructure of minority language processing in general and VML processing in particular. The next section will deploy the problem of Ede language processing in Vietnam, such as using Unicode font in Ede language text editor and converting Ede language documents using own fonts to Unicode font. The Final is the conclusion.

IV. USING UNICODE FONT IN THE EDE LANGUAGE DOCUMENTS

A. Ede language

Ede is an ethnic group in the national community of Vietnam. Currently, the population of the Ede ethnic group is over 330,000 people that have inhabited in the provinces of Vietnam such as DakLak, southern GiaLai, western PhuYen, and KhanhHoa. Ede language belongs to the Malayo-Polynesian language family and is related to many continental Austrone-sian languages [18].

TABLE II: EDE LANGUAGE CHARACTER SET

Consonant						Vowel					
B	b	Ƀ	ƀ	Č	č	A	a	Ă	a	Â	â
D	d	Đ	đ	G	g	E	e	Ĕ	ĕ	Ê	ê
H	h	J	j	K	k	Ê̆	ê̆	I	i	Ĭ	ĭ
L	l	M	m	N	n	O	o	Ŏ	ŏ	Ô	ô
Ñ	ñ	P	p	R	r	Ô̆	ô̆	Ơ	ơ	Ơ̆	ơ̆
S	s	T	t	W	w	U	u	Ŭ	ŭ	Ư	ư
Y	y					Ư̆	ư̆				

Ede language character set was formed in the later years of the XIX century. In 1851-1857 Latin characters were used to write Bibles by the Ede and XTieng languages for missionary purposes. This character set has been corrected many times and is called the Ede language character set [19], [20]. The Ede language character set is classified as Latin family, with 76 characters including uppercase and lowercase characters as in Table 2 [21], [22]. These characters are the basic components of almost all Unicode fonts and 8 characters (ê̆, ô̆, ơ̆, ư̆, Ê̆, Ô̆, Ơ̆, Ư̆) without in Unicode fonts.

B. Deploying solution

The solution using Unicode font in VML text is deployed in Ede language text editor the following steps:

- Step1: Ede language character set is split into three groups as Table 3, in which:
  - Group 1: 54 characters are in Vietnamese character set.
  - Group 2: 14 characters are not in Vietnamese character set but not in Unicode.
  - Group 3: 8 characters are neither in Unicode nor in the Vietnamese character set.

TABLE III: GROUPS OF EDE LANGUAGE CHARACTER SET

Group	Ede language character set							
1	A	a	Ă	ă	Â	â	E	e
	Ê	ê	I	i	O	o	Ô	ô
	Ơ	ơ	U	u	Ư	ư	B	b
	D	d	Đ	đ	G	g	H	h
	J	j	K	k	L	l	M	m
	N	n	P	p	R	r	S	s
	T	t	W	w	Y	y		
2	Ƀ	ƀ	Č	č	Ĕ	ĕ	Ĭ	ĭ
	Ñ	ñ	Ŏ	ŏ	Ŭ	ŭ		
3	Ê̆	ê̆	Ô̆	ô̆	Ơ̆	ơ̆	Ư̆	ư̆

- Step 2: The characters in group 3 are encoded in character and the diacritic combining '˘'. The example of combining characters in group 3 is shown in Fig. 2.

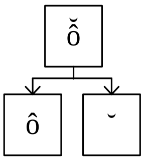


Fig. 2: The combination code example of a character in group 3

Characters in groups 2 and 3 are mapped to a subset of Unicode such as Latin supplement (H00A0:H00FF), Latin extended (H0100:H024F), Diacritical marks (H0300:H036F). The result of the mapping characters of groups 2 and 3 to Unicode is shown in Table 4.

TABLE IV: THE HEXADECIMAL VALUES OF CHARACTERS IN GROUP 2 AND 3

Group	Ede language character set			
2	Ƀ	ƀ	Č	č
	H0243	H0180	H010C	H010D
	Ĕ	ĕ	Ĭ	ĭ
	H0114	H0115	H012C	H012D
	Ñ	ñ	Ŏ	ŏ
	H00D1	H00F	H014E	H014F
	Ŭ	ŭ		
	H016C	H016D		
3	Ê̆	ê̆	Ô̆	ô̆
	H00CA	H00EA	H00D4	H00F4
	H0305	H0306	H0306	H0306
	Ơ̆	ơ̆	Ư̆	ư̆
	H01A0	H01A1	H016C	H016D
	H0306	H0306	H0306	H0306

- Step 3: Stipulating for typing method of characters in groups 2 and 3 is shown in Table 5.

TABLE V: TYPING METHOD OF CHARACTERS IN GROUP 2 AND GROUP 3

	Group 2	Group 3
Character	Ƀ ƀ Č č Ĕ ĕ Ĭ	Ê̆ ê̆ Ô̆ ô̆
Type	B~ b~ C^ c^ E^ e^ I^	Ê^ ê^ Ô^ ô^
Character	ĭ Ñ ñ Ŏ ŏ Ŭ ŭ	Ơ̆ ơ̆ Ư̆ ư̆
Type	i^ N~ n~ O^ o^ U^ u^	Ơ^ ơ^ Ư^ ư^

- Character '˜' is selected to replace the dash of characters (Ƀ, ƀ) and the tilde of characters (Ñ, ñ) in group 2.
- Character 'ˆ' is selected to replace the diacritic combining '˘' of characters in groups 2 and 3.

Two replacing characters are proposed aim help users to remember and visualize as Ede characters easily. Selecting two characters has also been checked on Vietnamese documents with over 1,000,000 words in the fields such as language, literature, history, social sciences, natural sciences, arts, sports. Result did not find combination characters in the checked texts.

- Step 4: The hexadecimal values and stipulated typing methods are installed on the WinVNKey application. This function interacts HTF file with the WinVNKey application for using WinVNKey typing tool in the text editor with Unicode font. A tool is proposed building to support the implementation of the above steps. It is named H&TES (Hexadecimal & Typing Ede Save).

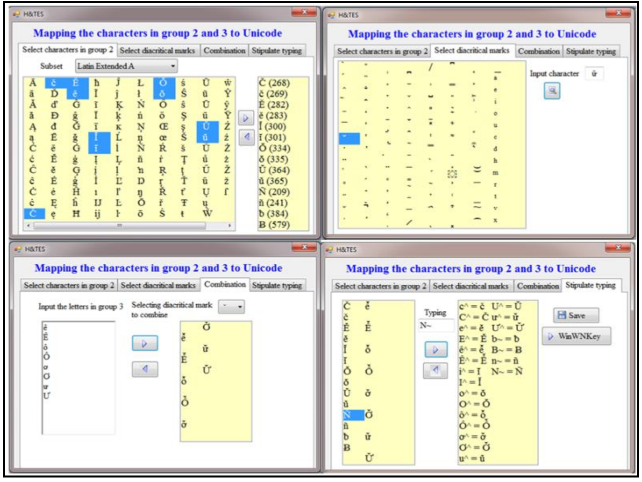


Fig. 3: The interfaces of functions in H&TES

The interfaces of functions in each Tab of H&TES are shown in Fig. 3 with the four functions:

1) Select character in group 2 of Tab 1 supports to select the characters in group 2 from the subset of Unicode. The characters of each subset in Unicode 2 display on the interface for selecting characters to map corresponding hexadecimal values.
2) Select diacritical marks of Tab 2 supports to select combining diacritical marks of characters in group 3.
3) Combination of Tab 3 supports to map of the characters in group 3 through inputting characters and selecting diacritical marks to map two corresponding hexadecimal values. Selected diacritical marks in Tab 2 will be items of the combo box in this Tab.
4) Stipulate for typing method of Tab 4 supports to input the typing method for characters in groups 2 and 3. This Tab has the following two functions:
    - Saving the hexadecimal values and typing methods of characters in groups 2 and 3 in a text file that is named HTF (Hexadecimal Typing File). The content of rows in the HTF file includes:
        - Stipulated typing method of the characters in group 2 and group 3;
        - Character '=' separating the stipulated characters and the mapped hexadecimal values;
        - Character '+' separating two hexadecimal values of characters in group 3.

The Illustration example of the rows in the HTF file is shown in Table 6.

TABLE 6: THE ILLUSTRATION ROWS IN HTF

Character	The content of rows
Ƀ	B~=H0243
Ơ̆	Ơ^=H01A0+H0306

- Interacting with the WinVNKey application will process data in the HTF file. The interactive function is implemented as Fig. 4

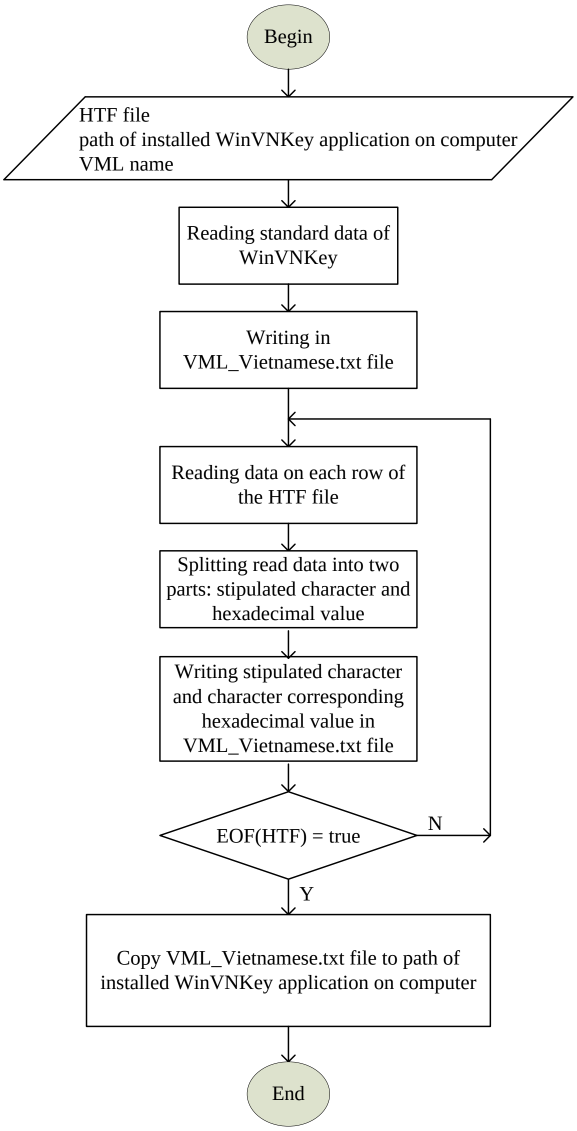

Begin
HTF file
path of installed WinVNKey application on computer
VML name
Reading standard data of WinVNKey
Writing in VML_Vietnamese.txt file
Reading data on each row of the HTF file
Splitting read data into two parts: stipulated character and hexadecimal value
Writing stipulated character and character corresponding hexadecimal value in VML_Vietnamese.txt file
EOF(HTF) = true
N
Y
Copy VML_Vietnamese.txt file to path of installed WinVNKey application on computer
End

Fig. 4: Interacting with the WinVNKey application

C. *Discussion*

Based on The solution using Unicode font in VML text to apply for Ede language processing, the paper has built the H&TES tool that has solved the problem of displaying Ede language character set in a multilingual environment with WinVNKey typing tool. Previous research results and research results of the paper are compared through the elements of displaying the Ede language character set with Unicode font. The evaluations by compared elements are shown in Table 7.

TABLE VII: COMPARING THE ELEMENTS TO DISPLAY EDE LANGUAGE CHARACTER SET

Element	TayNguyenKey	VnKey	H&TES
Unicode font	No use	No use	Used
Typing tool	Using Unikey, VietKey	Using VNKey	Using WinVNKey
Combination secondary key	12 keys	Characters as telex or number as VNI	2 keys: '~' and '^.'

The research result of TayNguyenKey fonts has integrated with the Vietnamese typing tool but still has to use its own fonts. Using own font set is the disadvantage of TayNguyenKey. Besides, TayNguyenKey has used 12 extra keys that are difficult to remember and not friendly for users.

VNKey research result has built its own typing tool and own fonts that can type Vietnamese and Ede language. Just as the result of the TayNguyenKey font set, VNKey typing tool has not used Unicode font and also has disadvantages when using its own font set.

The H&TES tool of paper overcame the disadvantages when using its own font set of the before research results.

Through H&TES, Ede language characters have been edited with Unicode fonts in text editor applications. The H&TES tool can also expand with the other ethnic languages that have Latin scripts.

V. CONVERTING THE EDE DOCUMENTS WITH OWN FONT TO UNICODE FONT

A. *Solution converting Ede documents*

Currently, Ede documents do not use Unicode font. Using own fonts in the text is an obstacle in digitizing to develop information systems. In order to contribute to solving the difficulties of using their own fonts in text, the paper proposed the solution for converting Ede documents with their own fonts to Unicode font. The model of the solution is shown in Fig. 5.

In this solution, the H&TES tool is also used for saving the typing method of the characters with their own font and the hexadecimal values of the Unicode encoded Ede language character set in the HTF file. This file is the input data of the text converting function. This function will search the characters of each row in the inputted file. They will be replaced with the Ede language character corresponding to the hexadecimal value in the input file.

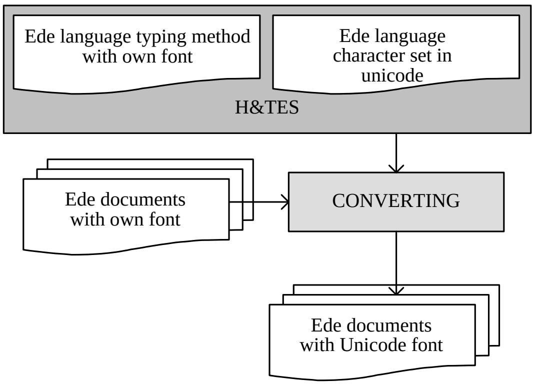

Ede language typing method with own font
Ede language character set in unicode
H&TES
Ede documents with own font
CONVERTING
Ede documents with Unicode font

Fig. 5: The model converting Ede documents with own font to Unicode font

From solution converting Ede language text with own font to Unicode font, a tool is proposed building and named CEDU. The converting function in CEDU will convert the Ede language text with own font in the clipboard or in file with extension as TXT, DOC, DOCX and RFT to Unicode. The converting function is implemented as Fig. 6.

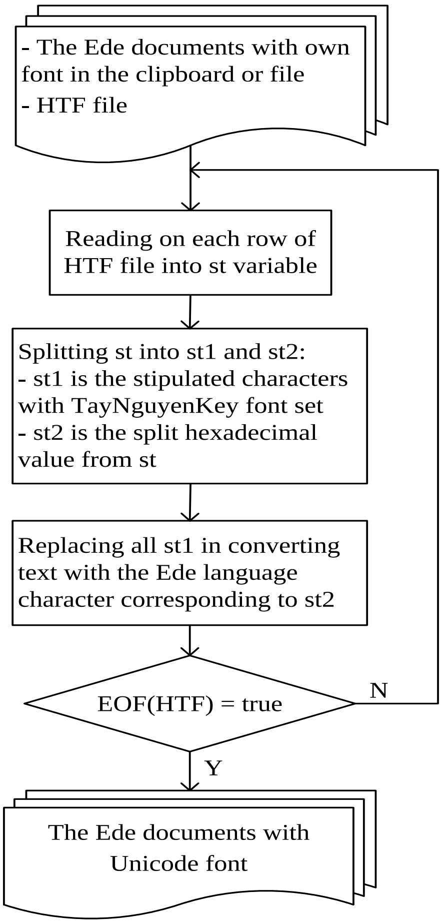

- The Ede documents with own font in the clipboard or file
- HTF file
Reading on each row of HTF file into st variable
Splitting st into st1 and st2:
- st1 is the stipulated characters with TayNguyenKey font set
- st2 is the split hexadecimal value from st
Replacing all st1 in converting text with the Ede language character corresponding to st2
EOF(HTF) = true
N
Y
The Ede documents with Unicode font

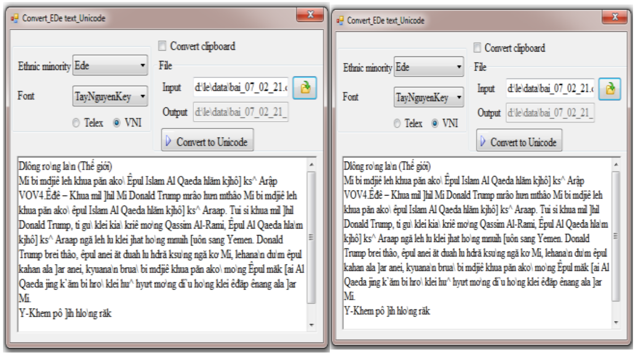
Fig. 6: The function of converting Ede documents with their own font to Unicode font

B. *Experimental results*

For evaluating the effectiveness of the converting function in the CEDU tool, the paper performed several experiments on input data sources from the bulletins of

Vietnam ethnic voice radio [23]. These have no misspelling and are checked. The font in the bulletins is TayNguyenKey font, VNI typing method.

Fig 7 illustrates the CEDU tool that converts TayNguyenKey font to Unicode font in an Ede language bulletin.



Dlông ro\ng la\n (Thế giới)
Mi bi mdjiê leh khua păn ako\ Êpul Islam Al Qaeda hlăm kjhô] ks^ Arập
VOV4.Êđê – Khua mil]hil Mi Donald Trump mrâo hưn mthâo Mi bi mdjiê leh khua păn ako\ êpul Islam Al Qaeda hlăm kjhô] ks^ Araap. Tui si khua mil]hil Donald Trump, ti gu\ klei kia\ kriê mo\ng Qassim Al-Rami, Êpul Al Qaeda hla\m kjhô] ks^ Araap ngă leh lu klei jhat ho\ng mnuih [uôn sang Yemen. Donald Trump brei thâo, êpul anei ăt duah lu hdră ksu\ng ngă kơ Mi, lehana\n du\m êpul kahan ala]ar anei, kyuana\n brua\ bi mdjiê khua păn ako\ mo\ng Êpul măk [ai Al Qaeda jing k`ăm bi hro\ klei hư^ hyưt mơ\ng di`u ho\ng klei êđăp ênang ala]ar Mi.
Y-Khem pô]ih hlo\ng răk

Dlông ro\ng la\n (Thế giới)
Mi bi mdjiê leh khua păn ako\ Êpul Islam Al Qaeda hlăm kjhô] ks^ Arập
VOV4.Êđê – Khua mil]hil Mi Donald Trump mrâo hưn mthâo Mi bi mdjiê leh khua păn ako\ êpul Islam Al Qaeda hlăm kjhô] ks^ Araap. Tui si khua mil]hil Donald Trump, ti gu\ klei kia\ kriê mo\ng Qassim Al-Rami, Êpul Al Qaeda hla\m kjhô] ks^ Araap ngă leh lu klei jhat ho\ng mnuih [uôn sang Yemen. Donald Trump brei thâo, êpul anei ăt duah lu hdră ksu\ng ngă kơ Mi, lehana\n du\m êpul kahan ala]ar anei, kyuana\n brua\ bi mdjiê khua păn ako\ mo\ng Êpul măk [ai Al Qaeda jing k`ăm bi hro\ klei hư^ hyưt mơ\ng di`u ho\ng klei êđăp ênang ala]ar Mi.
Y-Khem pô]ih hlo\ng răk

Fig. 7: CEDU tool

The experimental results, the CEDU tool is converted all typed characters with TayNguyenKey font to Unicode font in the 10 bulletins that are saved as type. DOC, .DOCX, RFT and . TXT (10 files per type) are shown in Table 8.

TABLE VIII: EXPERIMENTAL RESULTS OF CONVERTED TEXTS

Extension	Size	Words in text with TayNguyenKey font	Words in text with Unicode font
DOC	301 Kb	4,148	4,148
DOCX	168 Kb	4,148	4,148
RFT	433 Kb	4,148	4,148
TXT	53 Kb	4,148	4,148

Table 9 is the content of the HTF file created from the H&TES tool. The character column is VNI typing method of TayNguyenKey font. Content column is the hexadecimal values of the Unicode encoded Ede language characters.

TABLE IX: CONTENT OF HTF FILE

Character	Content	Character	Content
ă	a\ =H0103	Ñ	~ =H00D1
Ă	A\ =H0102	ñ	` =H00F1
ƀ	[=H0180	Ŭ	U| =H016C
Ƀ	{ =H0243	ŭ	u\ =H016D
č	] =H010D	ê̆	ê# =H00EA+H0306
Č	} =H010C	Ê̆	Ê$ =H00CA+H0306
ĕ	e\ =H0115	ô̆	ô# =H00F4+H0306
Ĕ	E| =H0114	Ô̆	Ô$ =H00D4+H0306
ĭ	^ =H012D	ơ̆	ơ\ =H01A1+H0306
Ĭ	& =H012C	Ơ̆	Ơ| =H01A0+H0306
Ŏ	O| =H014E	ư̆	ư\ =H01AF+H0306
ŏ	o\ =H014F	Ư̆	Ư| =H01B0+H0306

The results of converted texts by the CEDU tool have been manually tested on 10 original bulletins and found that the CEDU tool converted all typed characters with TayNguyenKey font to the Ede language characters with Unicode font.

C. Discussion

The CEDU tool has converted for files with extensions as TXT, DOC, DOCX, RTF and does not only limited to files with the. TXT or. RTF extension, like the converting function of Unikey and Unikey tools.

The CEDU tool has contributed to solve the disadvantages of exchanging Ede documents with their own font on the Internet or between computers, as well as reusing data sources in the research on Ede language processing.

VI. CONCLUSION

For contributing to digitize the VML documents for developing the VML information systems, the paper proposed the solution using Unicode font in VML text to edit VML text in applications. Paper deployed this solution in Ede language processing and implemented building the H&TES and CEDU tool.

The H&TES tool created the environment interacts into WinVNKey application that has been installed in the computer. This environment has contributed to solve the problem of using Unicode for Ede language text editing in a multilingual environment with the WinVNKey typing tool.

The CEDU tool has contributed to solve the disadvantages of exchanging Ede language texts with their own font and reusing data sources of the research results in Ede language processing.

ACKNOWLEDGMENT

This research is supported by Funds for Science and Technology Development of the Ministry of Education and Training under project number B2019-DNA-01.

Thank you for all authors of the papers quoted.

REFERENCES

[1] H.K. Phan, Building grammar to process text. Applying for Vietnam ethnic languages, in Proc. Anniversary of the Founding of the Vietnam Academy of Science and Technology, 1976-2006, (2006).

[2] P. Baker, A. Hardie, T. McEnery, and et al., 11A 67-Million Word Corpus of Indic Languages: Data Collection, Markup, and Harmonisatio, in: Proc. The LREC-Language Resources and Evaluation Conferences, (2002) 819–825.

[3] B. Williams, M.L. Forcada, K. Sarasola, 6th SaLTMiL Workshop on: Collaboration: interoperability between people in the creation of language resources for less-resourced languages, in: Proc. SALTMIL, Morocco, (2008).

[4] K. Sarasola, F.M. Tyers, M.L. Forcada, 7th SaLTMiL Workshop on: Creation and use of basic lexical resources for lessresourced languages, in: Proc. SALTMIL, Malta,(2010).

[5] M.L. Forcada, G.D Pauw, G.M. Schryver, K. Sarasola, F.M. Tyers, P.W. Wagacha, Language technology for normalisation of less resourced languages, in: Proc. SALTMIL, Turkey, (2012).

[6] S. Rob, The Unicode Standard, Mountain View Publishing, 2016.

[7] Unicode http://vi.wikipedia.org/wiki/Unicode. K.C. Le, Researching on Vietnam ethnic minority languages, Aboriginal Education World, Taiwan, 2013

[8] T.D. Tran, Researching on languages of ethnic minorities in Vietnam, Hanoi National University Publishing, (1999).

[9] K.Q. Truong, X.D. Tran, Cham language processing-Building English-Vietnamese-Cham multilingual text editing system¸Thesis of Information Technology Engineer, The University of Danang, (2003).

[10] Using the software of Nom, Thai and Cham language character sets, http://tintuc.hues.vn/dua-vao-ung-dung-phan-mem-chunom-chu-thai-chu-cham, (2012).

[11] H.T.M. Le, Building a Ede computer information processing system in text editor, Master thesis in Computer Science, (2002).

[12] D.K. Nguyen, TayNguyenKey - Supporting program for typing character set of ethnic minorities in the Central Highland, Dak Lak Department of Education, http://c3quangtrung.daklak.edu.vn/tai-nguyen/taynguyenkey-chuong-trinh-ho-tro-go-chu-cac-dan-toc-thieu-so-tay-nguyen.

[13] Q. Huy, N. Thuy, Bringing ethnic minority languages to VnKey's typing tool, https://www.vietnamplus.vn/dua-ngon-ngu-dan-toc-thieu-so-len-bo-go-vnkey/80904.vnp.

[14] S.Cam, The problem of Thai language font and typing tool, http://learntaidam.blogspot.com/2012/05/van-e-bo-go-va-fonttieng-thai.html, (2006).

[15] T.D. Tran, Researching on ethnic minority languages in Vietnam, Hanoi National University Publishing, (1999).

[16] T.B. Tran, Vietnamese & multilingual Keyboard Driver for Windows, http://winvnkey.sourceforge.net.

[17] T.M.L. Hoang, V. Souksan, H.K. Phan, Using Unicode in Encoding the Vietnamese Ethnic Minority Languages, Applying for the Ede Language, in. Proc. The International Conference on Knowledge and System Engineering, Hanoi,Vietnam, (2013) 137–148.

[18] Y Čang Niê Siêng, Y ČôČ Mlô, Ede language, DakLak Department of Education, (2007).

[19] Y.N. Kdam, Ede language book-1, Vietnam Education Publishing, (2013).

[20] Y.N. Kdam, Ede language book-2, Vietnam Education Publishing, (2013).

[21] Y.N. Kdam, Ede language book-3, Vietnam Education Publishing, (2013).

[22] DakLak Department of Education and Training, Ede language Grammar, Education Publishing, (2011).

[23] VOV4 ethnic radio system, Vietnames Broadcast, http://vov4.vov.vn/Ede.aspx.