

# Su Yi Language Translation Application

Zhou Hui<sup>1,\*</sup>, Qian Ye<sup>2</sup>

<sup>1</sup>School of Information Technology and Engineering Yuxi Normal University, 653100, China

<sup>2</sup>Yunnan Agricultural University, 650000, China

**Abstract.** Along with the mixed population, people lost their own dialect in the process of popularizing “Putonghua”. In order to protect this kind of traditional culture, more and more researchers have entered into the protection of endangered languages. But some of the current protection work is carried out manually, with consuming a lot of manpower and resources. In order to solve this problem, a dialect translation software is designed. The using of the software will automatically convert the traditional Yi language of ShanSu (Su Yi) to the international phonetic sign. It greatly improves the efficiency of the protection work..

## 1 Introduction

With the globalization of industry and commerce, the communication between people becomes more and more relaxed. But a serious problem behind the relaxed is the losing of the traditional culture, especially the traditional Chinese dialects. In the promotion of “Putonghua”, some young people, especially children, are developing the direction of “not to say” even “not understand”. So it is very important for the protection of the traditional dialect.

Yi language belongs to the Sino Tibetan language family, formed in six thousand years ago, popular in the ancient Shu. Su Yi is a kind of the Yi language. It mainly in the people of the ShanSu pass from mouth to mouth in Ershan county of Yuxi city in Yunnan province<sup>[1]</sup>. In order to protect the dialect, we mainly use artificial way to Su Yi manual translation for international standard phonogram. For the young people of Shansu to recognition, reading, memory and writing more faster, and recording native spelling system for language learning and transmission service. But this method needs to consume a lot of human resources, and the speed is slow. Therefore, in order to be able to complete the conversion of language more quickly and effectively, we have designed a translation software which can automatically convert Su Yi to the standard international phonetic alphabet, reduce the unnecessary loss of the artificial.

## 2 The conversion principle of Su Yi

In the process of translation in the Su Yi, it will use the same phonetic symbols if Su Yi language and Chinese have the same or similar pronunciation, and its unique language by adding letters or double letters to express. Tone symbols used the letters post standard method. A syllable is composed by consonants and vowels and tone<sup>[2]</sup>, as shown in Figure 1<sup>[1]</sup>.

Example			
	Su Yi	International Phonetic Alphabet	Meaning
Phrase	wa <sup>21</sup> lo <sup>21</sup>	waqlɔq	Network
	zi <sup>21</sup> ph <sup>21</sup>	yiavqpevq	Force
	ng <sup>33</sup> zi <sup>33</sup>	nahvri	Tears
Sentence	ŋa <sup>33</sup> dza <sup>33</sup> dza <sup>21</sup> ko <sup>55</sup> xa <sup>55</sup> .	Nga dza dzaq gol hal.	I've eaten.
	ni <sup>33</sup> khui <sup>33</sup> mi <sup>21</sup> ba <sup>33</sup> zi <sup>33</sup> zi <sup>33</sup> ko <sup>55</sup> ?	Nji kuimiq bba ri ri gol?	Have you ever been to Kunming?
	e <sup>33</sup> thu <sup>55</sup> ŋu <sup>55</sup> ma <sup>33</sup> ma <sup>21</sup> ŋa <sup>33</sup> .	Eh teelgeelma maq nga.	She is not Han.

**Figure 1.** The translation and implications of some phrases and sentences

Figure 1 shows the translation and implications of some phrases and sentences. The specific methods of translation and the translation of the results are presented in the following sections.

### 2.1 Initials

The consonant of Su Yi are 30. There is no complex consonant, only a single consonant. According to the place of articulation is divided into lips, teeth, gingival, palate consonant and so on. According to the pronunciation can be divided into plosive, affricate, laterals, etc. The consonant pronunciation and spelling are shown in Figure 2.

	Su Yi	International Phonetic Alphabet	Su Yi	International Phonetic Alphabet
Initials	p	b	th	t
	ph	P	d	dd
	b	bb	n	n
	m	m	ɬ	hl
	w	w	l	l
	f	f	te	j
	v	v	teh	q
	ts	z	dz	dj
	tsh	c	ɳ	nj
	dz	dz	ɛ	x
	s	s	z	y
	z	r	k	g
	t	d	kh	k
	g	gg	ŋ	ng
	x	h	y	hh

**Figure 2.** The consonant pronunciation and spelling

## 2.2 Vowel

The vowels of Su Yi are 28. There are 18 units and a pair retroflex vowel that are the opposite of the loose and tight vowels. There are 7 loose compound vowels, and in which only have one tight compound vowel. The vowel pronunciation and spelling are shown in Figure 3.

	Su Yi	International Phonetic Alphabet	Su Yi	International Phonetic Alphabet
Vowel	ɿ	i	e	eh
	ɛ	ah	a	a
	ə	e	u	ee
	u	u	o	o
	ɿ	iiv	i	iv
	ɛ	ehv	ɛ	ahv
	a	av	ə	ev
	u	eev	u	uv
	o	ov	ɛr	er
	ɛr	erv	ie	ieh
	ie	iah	ia	ia
	io	io	ui	ui
	ue	uah	ua	ua
	ia	iav		

Figure 3. The vowel pronunciation and spelling

## 2.3 Tone

The tone of Su Yi has 3: gaoping, medium, low. Respectfully mark by “l”, omitted, and “q”. Just like the shown in Figure 4.

	Su Yi	International Phonetic Alphabet	Example	
Tone	55	l	nɛ <sup>55</sup>	nahl
	33		nɛ <sup>33</sup>	nah
	21	q	nɛ <sup>21</sup>	nahq

Figure 4. The tone of Su Yi

## 2.4 Syllable

There are 7 types of Su Yi syllables: Vowels + tones, consonant + loose vowels + tone, consonant + tense vowels + tone, consonant + retroflex loose vowels + tone, consonant + retroflex tense vowels + tone, consonant + loose vowels + loose vowels + tone, and consonant + loose vowels + tense vowels<sup>[3]</sup>.

## 3. The overall design of Su Yi translation software

In order to convert Su Yi more faster, we designed a translation software, which can also translate the Su Yi without online. It can better solve the problem of artificial translation, improve the efficiency if translation work. Software mainly from four aspects to design, just like shown in Figure 5<sup>[4]</sup>.

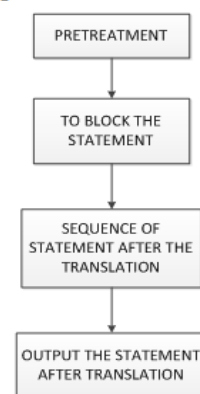


Figure 5. The design flow of the software

### 3.1 Pretreatment

The module mainly realizes four functions: to read the original sentence into the system; the original statement to block; save the location information of the block statements; standard processing of statements after the sub block.

The system of the original sentence is the process of acquiring the translation materials. The software will search, locate and acquire the original statement in the process. Ensure the original statement can be entered into the software according to the format and contents of the software.

Block is the most important part of the whole translation process. Whether the software can be correctly divided into blocks will directly determine whether the statement is normal or not. The first segmentation in preprocessing is mainly based on the punctuation in the sentence. At the same time, pay attention to the logical relationship of the original statement.

Software is required to store the position information of each block in the block, to determine the logical relationship of the block information, in order to avoid the confusion of word order after the translation.

After the statement is divided into blocks, each block of statements should be standardized. It will remove the secondary information that is no close relationship with the translation itself, and make sure the standard of each statement.

### 3.2 To block the statement

After the pretreatment into the key parts of the software: on the block statement translation. The second division will be at this time, according to the segmentation of initials, vowels and tones of the segmentation position, and save the position after each split. Then every letter of Su Yi will be accordance with the standards of translation for international phonogram.

### 3.3 Sequence of statements after the translation

After the translation, the location of the statement should be sorted out, and the logical relations should be

determined, and the corresponding punctuation should be added.

### 3.4 Output the statement after translation

The statement will be sorted in accordance with the format of the output display in front of the window.

## 4. The key technology of software

After the design of the software is mainly in the absence of the network and do not rely on the software on the computer, so the software after the completion of the design will be converted into application, in order to facilitate the use of direct on the computer. The software used in the design of the software is Visual Basic software<sup>[5]</sup>.

After the user enters the translation module, the user can directly operate in the corresponding dialog box. Translation interface as shown in Figure 6.

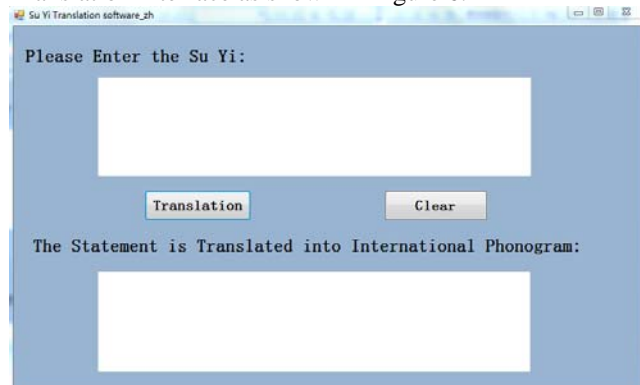


Figure 6. Translation interface

The key technology in the translation module is how to deal with the input statements. This software mainly uses the “two times block” to process the input information. The first block is carried out according to the punctuation mark, and the second sub block is carried on according to the syllable<sup>[6]</sup>.

## 5. The test of software

After design the software, it also has to testing the reliability of the software to determine whether the software can run reliably. This paper mainly uses the test method is the input test. First enter a period of abnormal data, such as Chinese characters, decimal, strings, etc., at this time the software should jump out of the information, the requirement to enter the correct statement. Just like the shown in Figure 7.

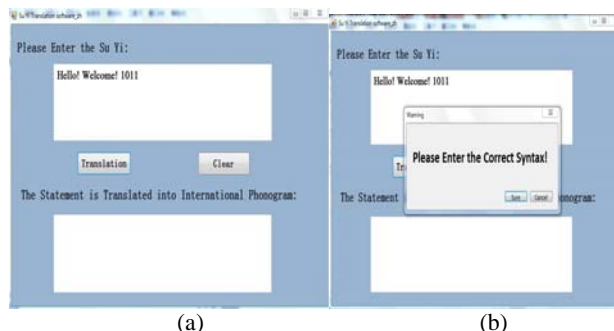


Figure 7. Abnormal Information

Figure 7 shows the following is the input of an abnormal data and the system to make the following tips. Through the figure, we can find the software can make a quick response when the input is not standard.

Second enter the short term, the corresponding dialog box should appear after the translation of the words. Just like the shown in Figure 8.

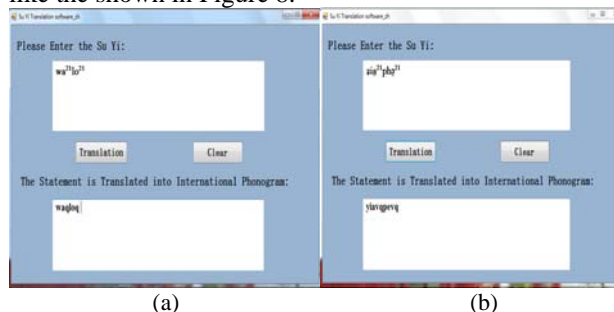


Figure 8. Short Term

Figure 8 shows the translation of the two different phrases respectively. Through the icon can be clearly seen when the input phrase, the software can achieve the right translation.

Through many tests, when the input is short, the software can complete the translation work.

Finally, enter the text of the great length. The quantity and quality of translation is an important index to evaluate the efficiency of translation software. Therefore, we have carried on the translation to the different length of the article. Just like the shown in Figure 9.

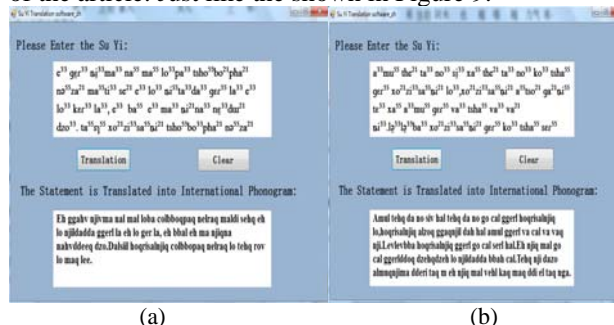


Figure 9. Paragraph Translation

Figure 9 shows the translation of two different articles. Through graphics can be seen, when we enter the different length of the article, the software is still able to complete the translation. In order to verify the reliability

of the software, we have carried out many experiments on different length of the article and the test results were drawn into a table, as shown in Table 1.

**Table 1.** The analysis of the accuracy of the translation of different lines.

The number of lines	< 20 lines	20~50 lines	50~80 lines	80~110 lines
The number of inputs	50	50	30	30
Average translation accuracy	100%	100%	99.8%	99.5%

Through the analysis of the table can be found, the software can be 100% complete the correct translation when less than 50 lines of the article. And the statement is only about 0.2% bit error rate after it is translated when more than 50 lines of the article. That is to say the software can still be very good to achieve the translation.

## 6. Concluding remarks

At this stage, with the use of the translation software, the original manual translation work has been greatly improved, the quantity and quality of the translation has been greatly improved. In the future, it will be considered to design the corresponding software for other ethnic minorities languages, and the integration of a variety of languages together, improve the efficiency of the protection of endangered languages.

## Acknowledgment

This paper was subsidized by the National Natural Science Foundation of China (61563055).

## References

- [1] Shi Changyan, Liu Yan. “*The Word Su Yi*”. Journal of Yuxi Normal University. 2014,30(10). PP:10-15.
- [2] Yu Wei, Zhou Ya, Wan Daili, Yang Xijing. “*Study on Yi isolated word recognition based on improved DTW*”. Journal of Kunming University of Science and Technology(Natural Science Edition). Vol.39 No.5 Oct.2014. PP:47-53.
- [3] Wang Li, Su Lianke, Tang Liming. “*Liangshan Yi adjective word formation marked and unmarked phenomenon*”. Journal of Southwest University for Nationalities (Humanities and Social Science). 2008,29(5). PP:81-85.
- [4] Luo Lixia. “*The Design and Implementation of Assembly Language to C Language Translation Software*”. Information Technology & Informatization. 2014,(10). PP:171-173.
- [5] Zhang Wenju. “*Using the Windows interface function*”. Journal of TaiYuan College of Economic Management Cadres. 2004,(1). PP:159-160.

- [6] Fang Lin, Fu Zheng, Zhang Yong. “*Small Translation Software for Mobile Phones*”. Business. 2012,(5). PP:60.