

A Technique for Speech Synthesis of Gurmukhi Embedded in LATEX Document

Jaswinder Singh¹, Tirath Singh², Sarabjit Singh³, Brahmjit Kaur⁴

¹Department of Computer Science, Guru Nanak Dev University, Amritsar, Punjab, India

²Computer Centre, Guru Nanak Dev University, Amritsar, Punjab, India

³Computer Section (Examinations), Guru Nanak Dev University, Amritsar, Punjab, India

⁴Department of Computer Science, Guru Nanak Dev University, Amritsar, Punjab, India

Abstract: *Text to speech conversion is one of the important applications of the natural language processing. Text to Speech application acts as an interface between text and speech to perform communication. Its main motive is to provide the globalization between two different parties by providing the effective communication. The work presented here, maps the Gurmukhi Script embedded in a Latex document using Pandey package to the corresponding Punjabi audio files. It is beneficial for the persons who want to perform the voice communication by using the presented material. In this work, parsing has been used so as to parse each and every word of the input text and then the next task was to identify the phonetic equivalent of the individual words. Once the phonetic equivalent has been obtained, the next work was to throw it to the audio player. At the final stage, the phonetic equivalent in accordance to the given input has been played to produce the person specific voice which is in Punjabi language. The whole work is performed on the JAVA platform.*

Keywords: Gurmukhi, Muharni, Latex, Speech Synthesis

1. Introduction

Text-To-Speech (TTS) synthesis system has a wide range of applications in everyday life. In order to make the computer systems more interactive and helpful to the users, especially physically and visibly impaired and illiterate masses, the TTS synthesis systems are in great demand for the Indian languages. Punjabi language is the regional language spoken by peoples in India, Pakistan and other countries with Punjabi migrants. The language is being written in Gurmukhi script in Indian Punjab, whereas in Shahmukhi script in Pakistani Punjab [10]. In this paper, TTS conversion of English text to corresponding Punjabi audio is performed. In this conversion, the concept of parsing is used. Parsing is the process of analyzing a string of symbols, either in natural language or in computer languages, according to the rules of a formal grammar. The term parsing comes from Latin *pars* (orationis), meaning part (of speech). To start the work, first of all, we have to record the Punjabi alphabet which has been recorded in mp3 format. Then we have to take the input text which is in the English language. Then we have to split each and every character from the text i.e. parsing is to be done. The next is to get the Punjabi phonetic equivalent to those characters. This whole process is done using JAVA platform. JAVA is being used because it provides portability, which means that computer programs written in the Java language must run similarly on any hardware or operating system platform. The goal of Java is to make the implementations of Java compatible. At the last stage, the phonetic equivalent which corresponds to the given text is to be connected and we get the final output in the Punjabi audio.

2. Related Work

Maninder Singh [5] performed work on „Text to Speech Synthesis for Numerals into Punjabi Language“. In this paper, he represents the work on text to speech synthesis for numerals in which speech waveform is generated and applied it to Punjabi speech synthesis using the general speech

synthesis architecture. In year 2006, Pardeep Gera [3] in „Text to Speech Synthesis for Punjabi Language“ has developed a TTS synthesizer for Punjabi language. In which synthesizer used an existing TTS named Dhvani that has been developed for Hindi and Kannada Language, as a model. Tapas Kumar have presented a method to design a text-to-speech conversion module using MatLab[13]. S. Gomathi and R.S. Umadevi performed a work on „Certain Approaches for a Complete Mathematical Text-To-Speech System“. This paper, discussed the teaching learning process of mathematical subjects made accessible to visually disabled students. It also aimed to provide users a friendly vocal interface with computer and allow visually impaired people to use the computer [12]. In year 2006, Nobuo Nukaga, Ryota Kamoshida in „Scalable Implementation of Unit Selection Based Text-to-Speech System for Embedded Solutions“ proposed two methods in order to implement unit selection-based text-to-speech engine into resource limited embedded systems [9]. Mirzanur Rahman and Sufal Das in „Parsing of part-of-speech tagged Assamese Texts“ produced a technique to check the grammatical structures of the sentences in Assamese text [8]. Richard Socher et al. have introduced recursive neural network architecture which has merged image segments or natural language words [11]. M.B. Chandak and Dr. R. V. Dharaskar in „Text to Speech Synthesis with Prosody feature: Implementation of Emotion in Speech Output using Forward Parsing“ have described various aspects of the design and implementation of speech synthesizer, which was capable of generating variable pitch output for the text. The concept of forward parsing was used to find out the emotion in the text and generate the output accordingly [6]. Jasleen Kaur & Gurpreet Singh Josan has addressed the issue of statistical machine transliteration from English to Punjabi [4]. Deepti Bhalla & others have done machine transliteration for English to Punjabi language pair using Rule-Based Approach [1]. D. Sasirekha and E.Chandra in „Text to Speech: A Simple Tutorial“ have made a clear and simple overview of working of text to speech system (TTS) in step by step process [2].

3. Proposed Work

The methodology followed is outlined as follows:

1. Parsing of the input text has been done and corresponding recorded sounds have to be fetched to produce corresponding Punjabi audio.
2. The sound of the Punjabi alphabets has been used as the basic unit for speech synthesis.
3. Speech database for the Punjabi language has been developed by recording the sounds of each and every alphabet.
4. Font character mapping of the input text has been done and then separated into Punjabi phoneme.
5. Phonemes have been searched in the file where audio for all alphabets are recorded and corresponding phonemes sounds have been concatenated to generate synthesized output speech.

3.1 Basic Flow of the Work

The basic flow of work for English text to Punjabi speech conversion starts as, first of all, the recording of the Punjabi alphabets is to be done. The chart of the Punjabi alphabets (ਮੁਹਾਰਨੀ) whose recording is to be done is shown below in the table 1:

By starting the work, get the input text in the English language written in the Latex file. LATEX package is used because it provides portability, Ease of collaboration, Free license and community support, Content before style, Reuse LATEX code. After getting the input text, perform the font character mapping to know that which letter corresponds to which character.

Table 1: Punjabi Alphabets with laga matra (ਮੁਹਾਰਨੀ)

ਮੁਹਾਰਨੀ											
ਅ	ਆ	ਇ	ਈ	ਉ	ਊ	ਏ	ਐ	ਓ	ਔ	ਅੰ	ਆਂ
ਸ	ਸਾ	ਸਿ	ਸੀ	ਸੁ	ਸੂ	ਸੇ	ਸੈ	ਸੋ	ਸੌ	ਸੰ	ਸਾਂ
ਹ	ਹਾ	ਹਿ	ਹੀ	ਹੁ	ਹੂ	ਹੇ	ਹੈ	ਹੋ	ਹੌ	ਹੰ	ਹਾਂ
ਕ	ਕਾ	ਕਿ	ਕੀ	ਕੁ	ਕੂ	ਕੇ	ਕੈ	ਕੋ	ਕੌ	ਕੰ	ਕਾਂ
ਖ	ਖਾ	ਖਿ	ਖੀ	ਖੁ	ਖੂ	ਖੇ	ਖੈ	ਖੋ	ਖੌ	ਖੰ	ਖਾਂ
ਗ	ਗਾ	ਗਿ	ਗੀ	ਗੁ	ਗੂ	ਗੇ	ਗੈ	ਗੋ	ਗੌ	ਗੰ	ਗਾਂ
ਘ	ਘਾ	ਘਿ	ਘੀ	ਘੁ	ਘੂ	ਘੇ	ਘੈ	ਘੋ	ਘੌ	ਘੰ	ਘਾਂ
ਚ	ਚਾ	ਚਿ	ਚੀ	ਚੁ	ਚੂ	ਚੇ	ਚੈ	ਚੋ	ਚੌ	ਚੰ	ਚਾਂ
ਛ	ਛਾ	ਛਿ	ਛੀ	ਛੁ	ਛੂ	ਛੇ	ਛੈ	ਛੋ	ਛੌ	ਛੰ	ਛਾਂ
ਜ	ਜਾ	ਜਿ	ਜੀ	ਜੁ	ਜੂ	ਜੇ	ਜੈ	ਜੋ	ਜੌ	ਜੰ	ਜਾਂ
ਝ	ਝਾ	ਝਿ	ਝੀ	ਝੁ	ਝੂ	ਝੇ	ਝੈ	ਝੋ	ਝੌ	ਝੰ	ਝਾਂ
ਟ	ਟਾ	ਟਿ	ਟੀ	ਟੁ	ਟੂ	ਟੇ	ਟੈ	ਟੋ	ਟੌ	ਟੰ	ਟਾਂ
ਠ	ਠਾ	ਠਿ	ਠੀ	ਠੁ	ਠੂ	ਠੇ	ਠੈ	ਠੋ	ਠੌ	ਠੰ	ਠਾਂ
ਡ	ਡਾ	ਡਿ	ਡੀ	ਡੁ	ਡੂ	ਡੇ	ਡੈ	ਡੋ	ਡੌ	ਡੰ	ਡਾਂ
ਢ	ਢਾ	ਢਿ	ਢੀ	ਢੁ	ਢੂ	ਢੇ	ਢੈ	ਢੋ	ਢੌ	ਢੰ	ਢਾਂ
ਣ	ਣਾ	ਣਿ	ਣੀ	ਣੁ	ਣੂ	ਣੇ	ਣੈ	ਣੋ	ਣੌ	ਣੰ	ਣਾਂ
ਤ	ਤਾ	ਤਿ	ਤੀ	ਤੁ	ਤੂ	ਤੇ	ਤੈ	ਤੋ	ਤੌ	ਤੰ	ਤਾਂ
ਥ	ਥਾ	ਥਿ	ਥੀ	ਥੁ	ਥੂ	ਥੇ	ਥੈ	ਥੋ	ਥੌ	ਥੰ	ਥਾਂ
ਦ	ਦਾ	ਦਿ	ਦੀ	ਦੁ	ਦੂ	ਦੇ	ਦੈ	ਦੋ	ਦੌ	ਦੰ	ਦਾਂ
ਧ	ਧਾ	ਧਿ	ਧੀ	ਧੁ	ਧੂ	ਧੇ	ਧੈ	ਧੋ	ਧੌ	ਧੰ	ਧਾਂ
ਨ	ਨਾ	ਨਿ	ਨੀ	ਨੁ	ਨੂ	ਨੇ	ਨੈ	ਨੋ	ਨੌ	ਨੰ	ਨਾਂ

ਪ	ਪਾ	ਪਿ	ਪੀ	ਪੁ	ਪੂ	ਪੇ	ਪੈ	ਪੋ	ਪੌ	ਪੰ	ਪਾਂ
ਫ	ਫਾ	ਫਿ	ਫੀ	ਫੁ	ਫੂ	ਫੇ	ਫੈ	ਫੋ	ਫੌ	ਫੰ	ਫਾਂ
ਬ	ਬਾ	ਬਿ	ਬੀ	ਬੁ	ਬੂ	ਬੇ	ਬੈ	ਬੋ	ਬੌ	ਬੰ	ਬਾਂ
ਭ	ਭਾ	ਭਿ	ਭੀ	ਭੁ	ਭੂ	ਭੇ	ਭੈ	ਭੋ	ਭੌ	ਭੰ	ਭਾਂ
ਮ	ਮਾ	ਮਿ	ਮੀ	ਮੁ	ਮੂ	ਮੇ	ਮੈ	ਮੋ	ਮੌ	ਮੰ	ਮਾਂ
ਯ	ਯਾ	ਯਿ	ਯੀ	ਯੁ	ਯੂ	ਯੇ	ਯੈ	ਯੋ	ਯੌ	ਯੰ	ਯਾਂ
ਰ	ਰਾ	ਰਿ	ਰੀ	ਰੁ	ਰੂ	ਰੇ	ਰੈ	ਰੋ	ਰੌ	ਰੰ	ਰਾਂ
ਲ	ਲਾ	ਲਿ	ਲੀ	ਲੁ	ਲੂ	ਲੇ	ਲੈ	ਲੋ	ਲੌ	ਲੰ	ਲਾਂ
ਵ	ਵਾ	ਵਿ	ਵੀ	ਵੁ	ਵੂ	ਵੇ	ਵੈ	ਵੋ	ਵੌ	ਵੰ	ਵਾਂ
ੜ	ੜਾ	ੜਿ	ੜੀ	ੜੁ	ੜੂ	ੜੇ	ੜੈ	ੜੋ	ੜੌ	ੜੰ	ੜਾਂ

The table from which the mapping has to be done is shown in Table 2. The top and bottom lines which is written in Punjabi language are used as an example to make the concept clearer.

Then the separation of the English words from the text is to be done. Once the text has been separated, then get the phonetic equivalent of the text. This process is carried out by performing parsing on each and every character.

Table 2: Keyboard mapping

A= ਾ	a= ਅ	1= ੧
B= ਭ	b= ਬ	2= ੨
C= ਚ	c= ਕ	3= ੩
D= ਡ	d= ਦ	4= ੪
E= ੇ	e= ਏ	5= ੫
F= ਫ	f= ਫ	6= ੬
G= ਘ	g= ਗ	7= ੭
H= ੀ	h= ਹ	8= ੮
I= ੀ	i= ਿ	9= ੯
J= ਝ	j= ਜ	0= ੦
K= ਖ	k= ਖ	'= '
L= ਲ	l= ਲ	= =
M= ਵ	m= ਮ	--= --
N= ਨ	n= ਨ	== = =
O= ੌ	o= ਓ	: = ੌ
P=	p= ਪ	' = ੌ
Q= ਢ	q= ੇ	" = "
R= ਰ	r= ਰ	, = ,
S= ਸ	s= ਸ	< = ੌ
T= ਠ	t= ਟ	> = ੌ
U= ੁ	u= ਊ	. = .
V= ਤ	v= ਵ	\ekonkarp= ੜ
W= ਥ	w=	\adk= ੱ
X=	x= ਛ	\tpp= ੰ
Y= ਧ	y= ਯ	\char126= ੷
Z= ਗ	z= ਜ	

And later, arrange these entire phonetic equivalents in a series respective to the given text. This whole procedure is carried out by using JAVA platform. NetBeans IDE 8.0 has to be installed & used on the system. JAVA has been used because of its characteristic which is portability, which means that computer programs written in the Java language must run similarly on any hardware or operating-system platform. The goal of Java is to make all implementations of Java compatible.

At last, we get the final output as the Punjabi audio. The diagram (Figure 1) showing the basic flow of the work is shown.

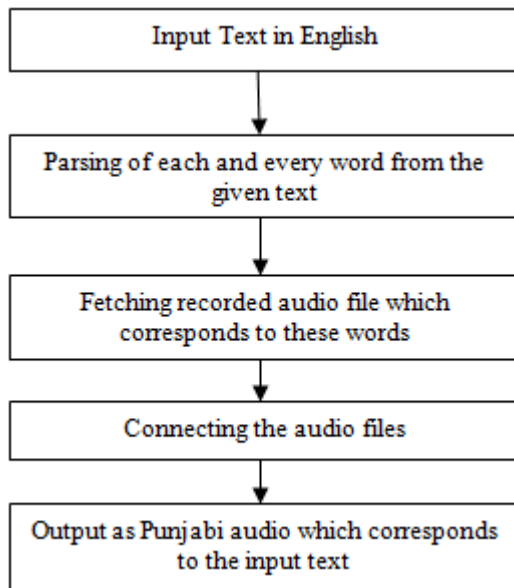


Figure 1: Flow of the proposed approach

4. Implementation

The various steps of implementation which have been followed in the conversion of English text into corresponding Punjabi audio are described below:

1. **Fetching the input string:** First of all, fetch the input text strings from the Latex document. After that split the input text in an individual character with laga matras.
2. **Storing string in array:** After splitting the input string, store the splitted string in an array. So that each and every character can get its individual position in an array.
3. **Finding length of string:** Then, find the length of the string so as to find the size of the string.
4. **Setting of pointer:** After finding the length of the string, set the pointer to the beginning of the array.
5. **Checking for different cases:** When the pointer is set, then check whether the pointer is less than the string length or not. If pointer is not less than the string length, fetch the audio file corresponding to that character where pointer points at. Then add in the playlist. But if, pointer is less than string length then check for the different cases like whether there is single character or double character. If two characters are there then, check whether there is any sign or symbol attached to it before or after. As we considered Punjabi language, in that case we have to look for the matra that whether it comes before the character or after the character. In this way, we get different cases.
6. **Adding identified sound to playlist:** After getting different cases, fetch the recorded audio file corresponding

to each and every character. Once we get the audio file of each character, and then add these in the playlist.

7. **Increment the pointer:** After that, increment the pointer to look for the next character. Then again, check for the different cases and after that add its corresponding audio file to the playlist.
8. **Playing the audio:** When we get different cases for the input string and fetched its corresponding audio files, then, the next step is to connect all the audio files. Add the fetched audio files to the playlist and play. We get the final output which is Punjabi audio corresponding to the given input text.

5. Results

The sample results of the proposed approach are shown below. The input English text written in Latex package is shown below in Figure 2. The conversion of the English text to Punjabi audio has done by performing parsing. The outputs of the individual words of the text are shown in Figure 2 which gives output in Punjabi audio.

```

{\gm gUrU nAnc dEv jI dA jnm 15 apr>l 1469
iv\adk C h\char126 ieaA}
{ਗੁਰੂ ਨਾਨਕ ਦੇਵ ਜੀ ਦਾ ਜਨਮ 15 ਅਪਰੈਲ 1469 ਵਿੱਚ
ਹੋਇਆ}
    
```

Figure 2: Input text in English language using Pandey in Latex Document

The output of the given text is shown below which gives output in Punjabi audio. The word “gUrU”, means “guru.” In the output, .mp3 represents silence after pronouncing one word i.e. Guru. The output of the second word is shown below in figure 4. The word “nAnc” means “Nanak”. The output for the third word is dEv jI is shown in figure 5. The word “dEv jI” means “Dev Ji.” The output for the word dA is shown in figure 6. The word “dA” means “Da.” The output for the word jnm is shown figure 7. The word “jnm” means “Janam.” The output for the word 15 is shown in figure 8. From the figure 9 it is quite clear that it is the output of the number 15. The output for the word apr>l is shown in figure 10. The word “apr>l” means “April.” The output for the number 1469 is shown figure 11. The output for the word “iv\adk C” is shown figure 12. The word “iv\adk C” means “Vich”. The word “h\char 126 ieaA” means “Hoyea.” In this way, we get the overall output of the English text into the Punjabi audio as shown in figure 13.

The overall output is, “Guru Nanak Dev Ji Da Janam 15 April 1469 Vich Hoyea.” as shown in figure 13.

```
run:
{\gUrU nAnc dEv jI dA jnm 15 apr>1 1469 iv\adk C h\char126
ieaA.}
Input String: gUrU nAnc dEv jI dA jnm 15 apr>1 1469 iv\adk C
h\char126 ieaA.
```

```
-----
CASE 5:
.mp3
-----
```

```
-----
CASE 1:
gU.mp3
-----
```

```
-----
CASE 1:
rU.mp3
-----
```

```
-----
CASE 5:
.mp3
-----
```

Figure 3: Audio output for the word gUrU (गुरु)

```
-----
CASE 1:
nA.mp3
-----
```

```
-----
CASE 5:
n.mp3
-----
```

```
-----
CASE 5:
c.mp3
CASE 5:
.mp3
-----
```

Figure 4: Audio output for the word

```
-----
CASE 1:
dE.mp3
-----
```

```
-----
CASE 5:
v.mp3
-----
```

```
-----
CASE 5:
.mp3
-----
```

```
-----
CASE 1:
jI.mp3
-----
```

```
-----
CASE 5:
.mp3
-----
```

Figure 5: Audio output for the word dEv

```
-----
CASE 1:
dA.mp3
-----
```

```
-----
CASE 5:
.mp3
-----
```

Figure 6: Audio output for the word

CASE 5:
j.mp3

CASE 5:
n.mp3

CASE 5:
m.mp3

Figure 7: Audio output for the word

CASE 5:
1.mp3

CASE 5:
5.mp3

CASE 5:
.mp3

Figure 8: Audio output for the word 15

CASE 5:
a.mp3

CASE 5:
p.mp3

CASE 1:
r>.mp3

CASE 5:
1.mp3

CASE 5:
.mp3

Figure 9: Audio output for the word

CASE 5:
1.mp3

CASE 5:
4.mp3

CASE 5:
6.mp3

CASE 5:
9.mp3

CASE 5:
.mp3

Figure 10: Audio output for the word

CASE 3:
ivadk.mp3

CASE 5:
.mp3

CASE 5:
C.mp3

CASE 5:
.mp3

Figure 11: Audio output for the word

CASE 1:
h\char126.mp3

CASE 6:
.mp3

CASE 4:
ie.mp3

CASE 2:
aA.mp3

CASE 5:
.mp3

Figure 12: Audio output for the word

ਗੁਰੂ ਨਾਨਕ ਦੇਵ ਜੀ ਦਾ ਜਨਮ 15 ਅਪਰੈਲ 1469 ਵਿੱਚ ਹੋਇਆ

Figure 13: Output in Punjabi

6. Conclusion and Future Work

In the presented work, parsing has been done through JAVA platform for converting Punjabi text embedded in Latex Document to Punjabi audio. TTS conversion software can have many applications in day to day life. It can be useful for blind people to read the document. Although these are some advantages of this presented work, but there are some limitations too. A person can understand a sentence only if it is pronounced correctly, but still some gaps in pronouncing need to be handled. So this application has come up with a better and user understandable pronunciation mechanism. The presented work methodology has been used for the Punjabi language and may also be used for other Indic language. An integration of a personal voice with the system can also be done. In the presented work, a sentence has been taken in which the individual digits have been mapped without any grouping and also the sentence is free from special symbols. So, in the future, input can be taken as a large vocabulary and with digits and special symbols.

References

- [1] Nisheeth Joshi, Deepti Bhalla and Iti Mathur. "Rule based transliteration scheme for English to Punjabi," International Journal on Natural Language Computing (IJNLC), 2013.
- [2] E.Chandra, D.Sasirekha. "Text to speech: A simple tutorial," International Journal of Soft Computing and Engineering (IJSCE), 2012.
- [3] Pardeep Gera. "Text-to-speech synthesis for punjabi language," M.Tech Thesis, 2006.
- [4] Jasleen kaur, Gurpreet Singh Josan. "Statistical approach to transliteration from english to Punjabi," International Journal on Computer Science and Engineering (IJCSE), 2011.
- [5] Karun Verma, Maninder Singh. "Text to speech synthesis for numerals into punjabi language," International Journal of Computational Linguistics and Natural Language Processing, 2013.
- [6] Dr. R.V.Dharaskar, Dr.V.M.Thakre and M.B.Chandak. "Text to speech synthesis with prosody feature: Implementation of emotion in speech output using forward parsing," International Journal of Computer Science and Security.
- [7] Sufal Das, Mirzanur Rahman and Utpal Sharma. "Parsing of part-of-speech tagged assamese texts," IJCSI: International Journal of Computer Science Issues, 2009.
- [8] Kenji Nagamatsu Nobuo Nukaga, Ryota Kamoshida and Yoshinori Kitahara. "Scalable implementation of unit selection based text-to-speech system for embedded solutions," IEEE, 2006.
- [9] Gurpreet Singh Lehal, Parminder Singh. "Text-to-speech synthesis system for punjabi language," Information Systems for Indian Languages Communications in Computer and Information Science, 139, 2011.
- [10] Andrew Y. Ng Richard Socher, Cliff Chiung-Yu Lin and Christopher D. Manning. "Parsing natural scenes and

natural language with recursive neural networks,” Proceedings of the 28th International Conference on Machine Learning, Bellevue, WA, USA, 2011.

- [11] R.S. Umadevi, S. Gomathi and Rohini. “Certain approaches for a complete mathematical text-to- speech system,” Indian Journal of Computer Science and Engineering (IJCSE) ISSN: 0976-5166, 2011.
- [12] Puspanjali Mohapatra, Tapas Kumar Patra, Biplab Patra. “Text to speech conversion with phonematic concatenation,” International Journal of Electronics Communication and Computer Technology (IJECCCT), 2012.

Author Profile



Jaswinder Singh is working as Assistant Professor at the Guru Nanak Dev University, Amritsar. He is having re Research experience in Wireless Sensor Networks, Language Processing and Neural Networks.



Tirath Singh has done Post Graduation from Punjabi University, Patiala and pursuing Ph.D. (Computer Science) at Guru Nanak Dev University, Amritsar. He is having four years experience working as DBA and Programmer in banking industry and is currently posted as System Manager at Guru Nanak Dev University, Amritsar.



Sarabjit Singh has done Post Graduation from Guru Nanak Dev University, Amritsar and is currently posted as System Manager at Guru Nanak Dev University, Amritsar.



Brahmjit Kaur has worked on the assignment as Post Graduate student at Guru Nanak Dev University, Amritsar Amritsar.