

A Step towards Making an Effective Text to speech Conversion System

Er.Sheilly Padda*, Er. Nidhi**, Ms. Rupinderdeep Kaur***

*(Department of Computer Science, Swami Vivekanand Institute of Engineering & Technology, Banur)

** (Department of Computer Science, Swami Vivekanand Institute of Engineering & Technology, Banur)

*** (Department of Computer Science, Thapar University, Patiala)

ABSTRACT

Text To Speech synthesis (TTS) is an online application to convert the text written in language to speech. This application helps user to type text and as output it gets the speech. The text inputted may be written by the operator or it may be scanned paper that is converted to speech. Though the user gets its output but still conversion of text to speech is not an easy task. It involves a great deal of work to be performed perfectly [1]. It is difficult to create database of number of phonemes, syllables, words for a particular language. And then concatenating these syllables, phonemes to produce an effective speech as output from the TTS engine. In this Paper we have discussed the text to speech conversion for Punjabi (Gurmukhi) language. The paper also discusses various issues which were found when converting text to speech. Various amendments can be done while converting Gurmukhi text to speech to solve the key issues occurred during conversion by already made converters.

Keywords: Amendments in proposed TTS, Gurmukhi, Text to speech conversion.

1. Introduction

Text to speech conversion involves Text-to-speech (T-T-S) systems based on the concatenation of speech units need a prosodic modification algorithm to adjust the prosodic features of the stored speech units to the desired output values. TTS application is not so easy as it requires more effort by the developer in cases text preprocessing where a text may be inputted in an ambiguous form by the different users, pronunciation problem where different words pronounced in different ways [5]. In this mostly homographs create much problem during pronunciation task. Lastly, the problem of prosody which includes the intonation, stress and duration which is the major challenging problem from many years.

2. Text to Speech Conversion

The text to speech conversion may be done in different steps: Text preprocessing, text analysis, text

phonetization, prosody generation and then the speech synthesis using various algorithms [7]. The steps followed to convert text to speech are described in Fig-1.

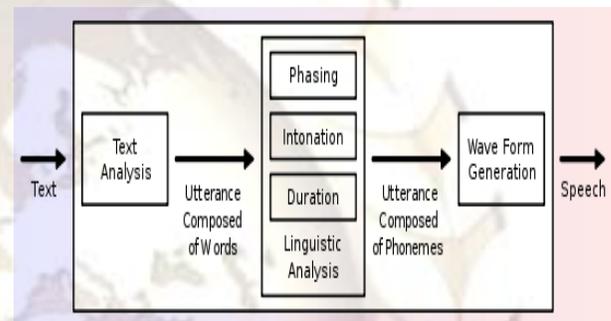


Fig 1: Steps in TTS system

2.1 Tokenization/Text preprocessing: Text preprocessing also called text normalization. A normal text contains many abbreviations, acronyms, special characters [2]. So they must be processed to convert them to segments called words or standard text format.

2.2. Syllabification

Combination of phonemes gives rise to next higher unit called syllables which is one of the most important units of a language [6]. A syllable must have a vowel called its *nucleus*, whereas presence of consonant is optional. In Punjabi seven types of syllables are recognized. These syllable types are: V, VC, CV, VCC, CVC, CCVC and CVCC; where V and C represent vowel and consonant respectively. Out of these seven types, occurrence of last two syllable types having sound clusters, is very rare in Punjabi.

2.3. Speech synthesis/Waveform production

The speech can be synthesized by concatenating different pieces of recorded speech from the database. Various methods can be employed for synthesized speech. Synthesis by reconstructing speech using parametric models such as the source-filter model for the speech production process. Acoustic frequency domain models or articulatory models can perform this

process. Speech can be synthesized using rule based, concatenation methods [4]. During speech synthesis various parameters must be considered for the speech like frequency, pitch, stress, intonation, noise levels. These parameters vary over time to create a wave of speech.

3. Gurmukhi

Gurmukhi derives its character set from old scripts of the Indian Sub-continent i.e. Landa (Script of North West), Sharda (script of Kashmir) and Takri (script of western Himalaya). It is a left-to-right syllabic script. A sentence illustrating Gurmukhi is given below:

ਖੁਰਬਾਨ ਜਿਓਨ ਮਨਕਾ ਨਾਮ

It has 38 consonants, 10 vowel characters, 9 vowel symbols, 2 symbols for nasal sounds and 1 symbol that duplicates the sound of a consonant.

3.1 Punjabi Phonemes

Phonemes are the abstract units and their pronunciation depends upon contextual effects, speaker characteristics and emotions [3].

Punjabi phonemes can be classified as: segmental phonemes and supra-segmental phonemes. Punjabi Language, like other Indian languages includes segmental phonemes, but not supra-segmental phonemes in its alphabet.

Segmental phonemes in Punjabi include twenty vowels and thirty eight consonants. Out of twenty vowels

(ੲ, ਈ, ਏ, ਐ, ਅ, ਆ, ਔ, ਉ, ਊ, ਓ) are non-nasalized

ten (ੲਿ, ਈਿ, ਏਿ, ਐਿ, ਅੰ, ਆੰ, ਔਿ, ਉੰ, ਊੰ, ਓਿ) are nasalized. And out of thirty eight consonants five

(ਙ, ਞ, ਣ, ਨ, ਮ) are nasalized and the remaining consonants are non-nasalized.

Punjabi vowels can be classified based on: opening of mouth, position of tongue tip and rounding of the tongue, whereas Punjabi consonants can be classified based upon: place of co-articulation and manner of articulation. Supra-segmental phonemes include stress, nasality, juncture, tone and intonation.

3.2 Gurmukhi script (ਗੁਰਮੁਖੀ)

Gurmukhi consists of vowels and consonants. The Gurmukhi consists of following vowels and vowel diacritics which are depicted in Fig-2.

ਅ	ਆ	ੲ	ਈ	ੳ	ਊ	ਏ	ਐ	ਓ	ਔ
a	ā	i	ī	u	ū	e	ai	o	au
[ə]	[ɑ]	[ɪ]	[i]	[ʊ]	[u]	[e]	[æ]	[o]	[ɔ]
ਕ	ਕਾ	ਕਿ	ਕੀ	ਕੁ	ਕੂ	ਕੇ	ਕੈ	ਕੇ	ਕੇ
ਕੰਨਾ	ਸਿਰੀ	ਬਿਰੀ	ਅੰਕੜ	ਦੁਲੰਕੜ	ਲਾਂਝਾਂ	ਦੁਲਾਂਝਾਂ	ਰੇੜਾ	ਕੰਨੜਾ	
kannā	sihārī	bihārī	aunkar	dulainkar	lānḥān	dulānḥān	hōṛā	kanaurā	
ka	kā	ki	kī	ku	kū	ke	kai	ko	kau

Fig-2: Vowels and Vowel diacritics (Laga Matra)

The consonants defined for Gurmukhi language are 38 in number and are described with their representation in Fig-3

ੳ	ੳਰ (ūrā)	ਅ	ਅੰਕੜ (ānrā)	ੲ	ੲਰੀ (irī)	ਸ	ਸੱਸਾ (sas'sā)	ਹ	ਹਾਰਾ (hāhā)
	u, ū, o		a, ā, ai, au		i, ī, e		sa [sə]		ha [hə]
ਕ	ਕੱਕਾ (kakkā)	ਖ	ਖੱਖਾ (khakhkhā)	ਗ	ਗੱਗਾ (gaggā)	ਘ	ਘੱਗਾ (ghaggā)	ਙ	ਙੱਙਾ (ñāñā)
	ka [kə]		kha [kʰə]		ga [gə]		gha [gʰə]		ṅa [ŋə]
ਚ	ਚੱਚਾ (caccā)	ਛ	ਛੱਛਾ (chachchā)	ਜ	ਜੱਜਾ (jajjā)	ਝ	ਝੱਝਾ (jhajjā)	ਞ	ਞੱਞਾ (ñāñā)
	ca [tʃə]		cha [tʃʰə]		ja [dʒə]		jha [dʒʰə]		ṅa [ɲə]
ਟ	ਟੈਂਕਾ (tainkā)	ਠ	ਠੱਠਾ (thaththā)	ਡ	ਡੱਡਾ (daddā)	ਢ	ਢੱਢਾ (dhaddā)	ਣ	ਣੈਂਕਾ (ṅāṅā)
	ta [tə]		tha [tʰə]		ḍa [d̪ə]		ḍha [d̪ʰə]		ṇa [ɳə]
ਤ	ਤੱਤਾ (tattā)	ਥ	ਥੱਥਾ (thaththā)	ਦ	ਦੱਦਾ (daddā)	ਧ	ਧੱਧਾ (dhaddā)	ਨ	ਨੈਂਨਾ (nannā)
	ta [tə]		tha [tʰə]		da [d̪ə]		dha [d̪ʰə]		na [nə]
ਪ	ਪੱਪਾ (pappā)	ਫ	ਫੱਫਾ (phaphphā)	ਬ	ਬੱਬਾ (babbbā)	ਭ	ਭੱਭਾ (bhabbā)	ਮ	ਮੈਂਮਾ (mamīmā)
	pa [pə]		pha [pʰə]		ba [b̪ə]		bha [b̪ʰə]		ma [mə]
ਯ	ਯੱਯਾ (yayyā)	ਰ	ਰਾਰਾ (rārā)	ਲ	ਲੱਲਾ (lallā)	ਵ	ਵੱਵਾ (vavvā)	ੜ	ੜਾਰਾ (rārā)
	ya [jə]		ra [rə]		la [lə]		va [və]		ṛa [r̪ə]
ਸ਼	ਸੱਸਾ (śasśā)	ਖ਼	ਖੱਖਾ (khakhkhā)	ਗ਼	ਗੱਗਾ (gaggā)				
	śa [ʃə]		kḥa [kʰə]		gḥa [gʰə]				
ਜ਼	ਜੱਜਾ (zazzā)	ਫ਼	ਫੱਫਾ (faffā)	ਲ਼	ਲੱਲਾ (lallā)				
	za [zə]		fā [fə]		lā [lə]				

Fig-3: Consonants (Vianjans)

Gurmukhi also provides the facility to write different numerals in Punjabi and here in Fig-4 a list of numerals representation in Punjabi language is shown .

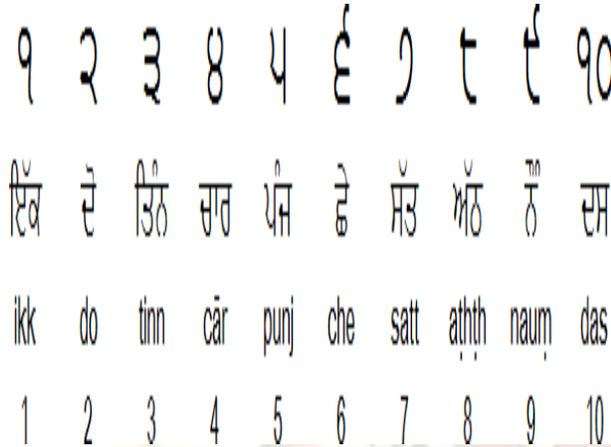


Fig-4: Description of Numerals in Gurmukhi

4. Amendments: A step towards modifying TTS

TTS is a text to speech system and is used to produce clear speech. But still, it suffers from lack of some important features in it. These features were thoroughly studied so that they can be solved in our proposed system. The various key issues which were seen in text to speech convertor for English language using some online software were: Special characters not were recognized. The problem of not recognizing the symbols like dot (.), question mark (?), hash (#) has also been noted. The database contains only few recorded voices to be used for speech synthesis. The current software leads to different pronunciation of a particular text. The speaker's pitch, stress of uttering text varied according to age group. Word limit is also very less for inputting text to be converted into speech. Text normalization was another problem in TTS convertor. So, these issues or the problem may be avoided in our proposed TTS system for Punjabi language. We will put effort and work out for improving our TTS convertor by taking into consideration following points.

- Increase the word limits of entering text.
- We can increase the database by recording no. of voices with different pitch and different frequencies.
- We can make our database effective by so that it requires less time to execute and saves our time and effort.

- We may allow recognition of numerals, abbreviations, acronyms, homographs in our TTS.
- Allowing the file to be saved as in format like .wav, .mp3 to be played in car, portable devices.

5. Conclusion

Using the above said points in amendments we can improve our text to speech convertor and can improve the naturalness of TTS. The Text to speech conversion may seem effective and efficient to its users if it produces natural speech and by making several modifications to it. Many of the above said issues in TTS system have been identified using some of the softwares for simple English language. These issues are been solved in our proposed system of text to speech conversion for Punjabi (Gurmukhi) language. We would create a system that would be able to solve problems of various individuals in their busy life and specially for the people with low vision or reading disabilities as it would help them to listen to their emails while relaxing, listen eBooks, study for exams by listening to notes.

References

- Nicolas Léwy and Thomas Hornstein, *Text-to-Speech Technology: A Survey of German Speech Synthesis Systems*.
- Michael H. O'Malley and Berkeley Speech Technologies, *Text To Speech Conversion Technology*.
- Parminder Singh and Gurpreet Singh Lehal, *Text-To-Speech Synthesis System for Punjabi Language*.
- Yousif A.El-Imam and Karima Banat, *Text to Speech conversion on Personal Computer*.
- Eduardo R. Banga and Carmen Garcia-Mateo, *Shape-invariant pitch- synchronous text to speech conversion*.
- M. Nageshwara Rao, Samuel Thomas, T. Nagarajan, and Hema A. Murthy, *Text-to-Speech Synthesis using syllable-like units*.
- Suresh Kumar Thakur and K.J. Satao, *Study of Various kinds of Speech Synthesizer Technologies and Expression for Expressive Text To Speech Conversion System*.