Text To Speech for Marathi Language using Transcriptions Theory

Sangramsing Nathusing Kayte
Research Scholar
Department of Computer Science & IT
Dr. Babasaheb Ambedkar Marathwada University,
Aurangabad

ABSTRACT
This research paper presents the approach towards converting text to speech using new methodology. The text to speech conversion system enables user to enter text in Marathi and as output it gets sound. The paper presents the steps followed for converting text to speech for Marathi language and the algorithm used for it. The focus of this paper is based on the tokenisation process and the orthographic representation of the text that shows the mapping of letter to sound using the description of language's phonetics. Here the main focus is on the text to IPA transcription concept. It is in fact, a system that translates text to IPA transcription which is the primary stage for text to speech conversion. The whole procedure for converting text to speech involves a great deal of time as it's not an easy task and requires efforts.

Keywords
IPA, Orthographic representation, phonetic, tokenization, transcription.

1. INTRODUCTION
The In today's life everyone wants that the present computer systems should behave like humans and proved to be user friendly. Even many of the great research scholars had dreamed of involving the machines in every facet of human life. With the growth of the power of computing machines their applications in modern day to day life is also rising [1]. This paper depicts one of the modern technologies that is the technology of text to speech convertor [2]. This paper presents the frame work of text to speech conversion methodology. The text to speech system has many applications that are useful in our life such as for video conferencing, helping people with reading difficulty, for visual aids process, Elementary and Applied Research, Telecommunications and Multimedia, Educational State, Screen reader programs for the visually impaired [3][4].

In this paper we have presented the linguistic features of Marathi text stating its phonetic representation. Through the knowledge of language's phonological orthographic we can develop a more promising text to speech conversion [4][5]. The first step in our text to speech conversion task is having the text tokenisation and to determine the token to sound concept.

2. OVERVIEW OF TEXT TO SPEECH SYSTEM
The text to speech conversion is carried out with the processing of text normalization (text preprocessing), text analysis, syllabification, prosody generation and finally the speech production [6][7][15]. The process of text analysis, text normalization and prosody generation all come under natural language processing while the speech generation is a part of digital processing system. A general view of the system is depicted in the Fig.1
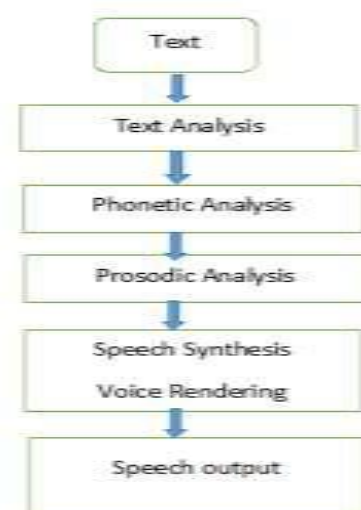

Fig 1: Overview of TTS [8]

The text to speech conversion involves a number of steps that are to be followed. The first step includes Tokenisation/Text preprocessing where the inputted text is broken down into smallest unit i.e. tokens. Tokenisation also includes the expansion of abbreviations.

The second step involves Syllabification process includes the combination of the small units to represent them as phonemes that identify the sound associated with each unit and we also called them as syllables.

In Marathi seven types of syllables are recognized [9][10].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 Marathi.

The third step is Speech Synthesizer which involves the generation of speech as output. There are various techniques that can be used for synthesizing speech [10]. The speech can be produced by concatenating different pieces of recorded speech from the database. Also, speech can be produced from concatenation of phonemes, diphones, and unit selection techniques and by Formant synthesis where the synthesized speech output is created using additive synthesis and an acoustic model (physical modelling synthesis).Parameters such as fundamental frequency, voicing, and noise levels are varied over time to create a waveform of artificial speech.

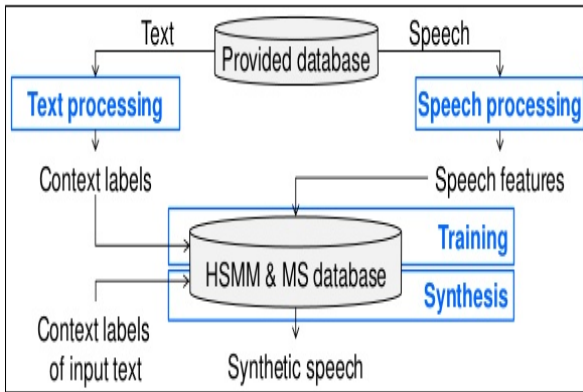## 3. ARCHITECTURE AND IMPLEMENTATION

Fig-2: Architecture of TTS

The entire process of converting text to speech includes analysis and synthesis [11] [16-19]. Although analysis part is similar for every techniques used for text to speech conversion, the difference is only in the synthesis part. The analysis part deals with determining the orthographic description of a particular language.

The implementation of text to speech conversion goes through various steps as depicted in Fig-2. The first step identifies the text for the processing. The text normalization is carried out. After the text has been normalized, the syllabification process is carried out where the input text is broken down into smallest unit called syllables

3.1 Text to IPA transcription

The phonetization is done which is the process of representing vocal sounds by signs and written characters. There are number of phonetic sounds that represent a letter of particular language. These phonetic sounds along with their written representation can be combined to generate sound using various speech synthesis tools.

The orthographic representation of text enables to represent the syllables with phonetic representation [12]. The orthographic set for a language consists of letter to sound rules. These rules define the function mapping of sequence of letter to sound segments. Specifically, the orthographic representation determined for a language, specifies the initial stages for text to speech conversion [13].The phonetic notation can be represented for Marathi language using an IPA chart (International Phonetic Alphabet).The IPA chart defines is an alphabetic system of phonetic notation based on some Latin alphabets [14]. They can be used for representation of sounds of spoken language [16-19].

The principal vowels for Marathi can be depicted from Fig-3. The principal vowels are symmetrically distributed on a standard vowel chart: three front vowels, two central vowels, and three back vowels. The three back vowels are rounded [14].
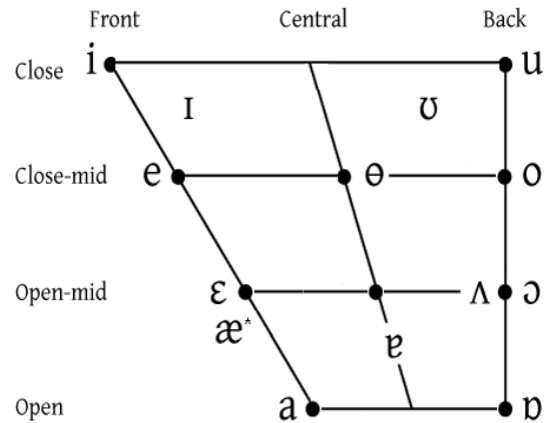

Fig-3: Vowel chart

B) Consonants IPA chart

Representations of consonant sounds outside of the core set are created by adding diacritics to letters with similar sound values.Fig-4 represent the consonants and their IPA symbols[1][3] [9].

Fig-4: IPA chart for consonants

The methodology followed here translates the text to IPA transcription using some defined rules. These rules are used for determining the letter orthographic features. These orthographic features can further be used by the speech synthesis tool to generate sound[4][16-19].

4. CONCLUSION

In this paper, we have described how the phonological attributes of Marathi orthography may be used to develop a text-to-speech system. The phonetic nature of the language was surveyed, and the set of rules for deriving these sounds using the orthography were made explicit and carefully deconstructed. Although the work done is not complete as it has been done until determining the linguistic features of Marathi text while the final signal processing aspect of the task are un explored. In the near future we would love to complete the text-to-speech system, the text to speech system that is able to be useful to the people in need.

5. REFERENCES

[1] Sangramsing Kayte, Monica Mundada "Study of Marathi Phones for Synthesis of Marathi Speech from Text" International Journal of Emerging Research in Management &Technology ISSN: 2278-9359 (Volume-4, Issue-10) October 2015

[2] Sangramsing N.kayte "Marathi Isolated-Word Automatic Speech Recognition System based on Vector Quantization (VQ) approach" 101th Indian Science Congress Jammu University 03th Feb to 07 Feb 2014

[3] Sangramsing Kayte, Monica Mundada, Dr. Charansing Kayte "Di-phone-Based Concatenative Speech Synthesis System for Hindi" International Journal of Advanced Research in Computer Science and Software Engineering -Volume 5, Issue 10, October-2015

[4] Sangramsing Kayte, Monica Mundada, Dr. Charansing Kayte "Di-phone-Based Concatenative Speech Synthesis Systems for Marathi Language" OSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 5, Ver. I (Sep –Oct. 2015), PP 76-81e-ISSN: 2319 –4200, p-ISSN No. : 2319 –4197 www.iosrjournals.org

[5] Monica Mundada, Sangramsing Kayte, Dr. Bharti Gawali "Classification of Fluent and Dysfluent Speech Using KNN Classifier" International Journal of Advanced Research in Computer Science and Software Engineering Volume 4, Issue 9, September 2014

[6] Simon King, "An introduction to statistical parametric speech synthesis", Indian Academy of Sciences, October 2011, pp. 837–852.

[7] Sangramsing Kayte, Monica Mundada, Dr. Charansing Kayte " Performance Calculation of Speech Synthesis Methods for Hindi language IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 6, Ver. I (Nov -Dec. 2015), PP 13-19e-ISSN: 2319 –4200, p-ISSN No. : 2319 –4197

[8] Sangramsing Kayte, Dr. Bharti Gawali "Marathi Speech Synthesis: A review" International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 6 3708 – 3711

[9] Sangramsing Kayte, Monica Mundada, Dr. Charansing Kayte "A Corpus-Based Concatenative Speech Synthesis System for Marathi" IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 6, Ver. I (Nov -Dec. 2015), PP 20-26e-ISSN: 2319 –4200, p-ISSN No. : 2319 –4197

[10] Sangramsing Kayte, Monica Mundada, Dr. Charansing Kayte "A Marathi Hidden-Markov Model Based Speech Synthesis System" IOSR Journal of VLSI and Signal

Processing (IOSR-JVSP) Volume 5, Issue 6, Ver. I (Nov -Dec. 2015), PP 34-39e-ISSN: 2319 –4200, p-ISSN No. : 2319 –4197

[11] Dhvani-TTS System for Indian Languages,(http://dhvani.sourceforge.net), 2001.

[12] Prof. Ian Maddieson,"Applied Phonetics: Portuguese Text-to-Speech", University of California, Berkeley Linguistics 110: May 16, 2003.

[13] Sangramsing Kayte, Monica Mundada, Dr. Charansing Kayte "Implementation of Marathi Language Speech Databases for Large Dictionary" IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 6, Ver. I (Nov -Dec. 2015), PP 40-45e-ISSN: 2319 –4200, p-ISSN No. : 2319 –4197

[14] International Phonetic Association (IPA), Handbook

[15] Sangramsing Kayte, Monica Mundada, Santosh Gaikwad, Bharti Gawali "PERFORMANCE EVALUATION OF SPEECH SYNTHESIS TECHNIQUES FOR ENGLISH LANGUAGE " International Congress on Information and Communication Technology 9-10 October, 2015

[16] Sangramsing Kayte, Monica Mundada,Dr. Charansing Kayte" Speech Synthesis System for Marathi Accent using FESTVOX" International Journal of Computer Applications (0975 – 8887) Volume 130 – No.6, November2015

[17] Sangramsing Kayte, Monica Mundada,Dr. Charansing Kayte "Screen Readers for Linux and Windows – Concatenation Methods and Unit Selection based Marathi Text to Speech System" International Journal of Computer Applications (0975 – 8887) Volume 130 – No.14, November 2015

[18] Sangramsing Kayte, Monica Mundada,Dr. Charansing Kayte " Performance Evaluation of Speech Synthesis Techniques for Marathi Language " International Journal of Computer Applications (0975 – 8887) Volume 130 – No.3, November 2015

[19] Sangramsing Kayte, Monica Mundada, Jayesh Gujrathi, " Hidden Markov Model based Speech Synthesis: A Review" International Journal of Computer Applications (0975 – 8887) Volume 130 – No.3, November 2015