

## **Research and Development of various Text to Speech synthesisers in Gujarati Language**

Vishal P Tank<sup>1</sup>, Sarman K Hadia<sup>2</sup>, Manthan S Manavadaria<sup>3</sup>

<sup>1</sup>V T Patel Department of Electronics & Communication Engineering, CSPIT & Charotar University of Science & Technology, Email: vishaltank.ec@charusat.ac.in

<sup>2</sup>V T Patel Department of Electronics & Communication Engineering, CSPIT & Charotar University of Science & Technology, Email: skhadia.ec@charusat.ac.in

<sup>3</sup>V T Patel Department of Electronics & Communication Engineering, CSPIT & Charotar University of Science & Technology, Email: manthanmanavadaria.ec@charusat.ac.in

**Abstract**— During last couple of years as the speech and language processing has provide it's significant in various fields. Text to speech conversion (Synthesiser) is very important aspect of speech and language processing field. Researcher has provided huge contribution to develop text to speech synthesiser in many foreign languages English, German etc. In India text to speech synthesiser is developed in many languages out of 22 official local Indian languages. This paper primary focus on various work done for text to speech synthesiser developed in Gujarati Language.

**Keywords**— Text to speech synthesiser, Indian Gujarati Language, Speech quality, Speech corpus

### I. INTRODUCTION

The meaning and purpose of synthesiser of speech is to convert the text to the synthetic speech (TTS- Text to Speech) which is as close to real speech as possible in acquiescence with the pronunciation touchstone of special language [1]. The objective of TTS system is to read electronic texts in the form of a book, and generate synthesis speech as shown in figure-1. In the current era of fast paced technology development in the field of information and communication technology such synthesisers can be used in various applications like artificial intelligence, robotics, in information referral systems, it can also be useful to help blind people or who lost reading ability [2,3].

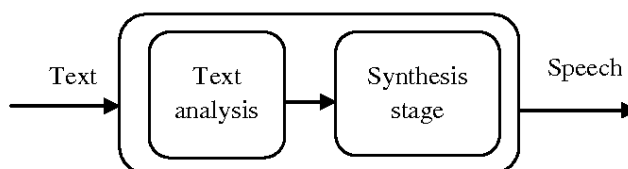


Fig. 1 General Process of any TTS system

Primary factors like its similarity to normal human speech (naturalness) and its intelligibility (ease of understanding by the listener) defines the quality of a speech synthesizer. Ideally, a speech synthesizer should be both natural and intelligible, and speech synthesis systems always attempt to maximize both the characteristics [4].

Further this Paper is organised as per 2) Requirement of TTS in Indian Languages, 3) Types of Synthesiser 4) Review of work done for Text to speech Synthesiser in Gujarati languages 5) Conclusion.

## II. REQUIREMENT OF TTS IN INDIAN LANGUAGES

India is Asian country which holds world's second highest numbers of languages are spoken. According to census of India 2001, 122 major languages and 1599 other languages are spoken in India these counts may vary from one sources to another because of dialect and language defines. As on 1st Dec 2007, The Eighth Schedule of the Constitution of India, 22 regional languages were approved based on speaking population as per following different family 1) Indo Aryan Family (Assamese, Bengali, Dogri, Gujarati, Hindi, Punjabi, Sindhi, Konkani, Maithili, Marathi, Nepali, Odia, Urdu, Kashmiri, Sanskrit) 2) Dravidian Family (Kannada, Telugu, Tamil, Malayalam,) 3) Sino Tibetan Family (Bodo, Meitei) 4) Austroasiatic Family (Santali) [5].

Gujarati is an official national and regional language of India. Gujarati is one of the important western Indian languages. Gujarati has many dialects i.e. Kathiyawadi, Sorathi, Surti, Charotari etc [6]. It is spoken by approximately 55 million people, making it the twenty-third most widely spoken language in the world today. Gujarati is spoken not only in India but also in other countries like USA, Tanzania, Uganda, Pakistan, and Kenya. With this understanding the importance of developing an TTS system for Gujarati language is highly recommended and hence review of it is not just logical but heartfelt.

Various Speech synthesiser has been developed in regional Indian language like Tamil [7], Malayalam [8], Hindi [9] etc.

## III. TYPES OF TTS SYNTHESISER

Symbolic prosody data is used by the synthesizer to generate speech using a specific method. There are three main categories of speech synthesis techniques. i. Articulator synthesis ii. Formant synthesis iii. Concatenative synthesis

**Articulator synthesis:** Articulator synthesis aims to simulate computationally the neurophysiology and biometrics of speech production. Articulator synthesis uses mechanical and acoustic model of speech production to synthesize speech. This synthesis produce intelligible synthetic speech, but its output is still far from natural sound [10, 11].

**Formant synthesis:** In this system representation of individual speech segments are stored on a parametric basis. The parametric are those of low level Holmes formant synthesizer for each segment. There is a single value for each parameter. This means single acoustic segment speech [10].

**Concatenative synthesis:** Concatenative synthesis is a technique for synthesizing sound by Concatenative samples of recorded sound called units. The duration of units is not strictly defined and may vary according to the implementation, roughly in the range of 10 ms up to 10 seconds. It is used in speech synthesis to generate user specific sequence of sound from a database built from recording of other sequences [11]. This methodology has the advantage in its simplicity and easy to implement [12].

## IV. REVIEW OF WORK DONE FOR TEXT TO SPEECH SYNTHESISER IN GUJARATI LANGUAGE

JJ Kothari et al. have built Gujarati Synthesiser based on Concatenative synthesis type. Author have Initially stored total 872 phonemes (34 Half Constants, 10 Digits, 35 (34 Constants+"V")\*12 Anuswar, 10 Digits, 34 Consonants\*12 Diacritic) and they were consolidated using Free audio editor software 2015. One of the factors regarding success of a good TTS system using concatenative method is depend upon how speech database is organized. Here authors have construct speech database in such a way that speech units are associated with phoneme and its ASCII code. Author found two benefits by applying this methodology: 1] phoneme matching and speech unit searching becomes very easy. 2] As Indian languages are phonetic in nature, one can easily apply same methodology for other Indian languages like Marathi, Bengali and Hindi etc [13].

Jashin Marina at al. have develop a Gujarati Synthesiser based on Concatenative synthesis for Gujarati language. Authors have stored 426 phonemes including vowels and consonant and their combination. The combination include here is C, CV, V, VC where C is the consonant and V is the vowel. The entire 426 phonemes database is created in the form of wave files (.wav) by using PRAAT software and Fs is 44100 Hz Where 13 phonemes are vowels and 34 phonemes are consonant and rest are their combinations. Perceptual evaluation of speech quality(PESQ) is obtained in MATLAB software which range from 3.36 to 4.14 (5 means Very natural, no degradation 4 means Fairly natural , little degradation 3 means Somewhat natural, somewhat degraded) [14].

Swati Talesara et al. have we have developed an algorithm and large speech corpus based on automate labeling of at syllable-level and build corresponding TTS in Gujarati language. The segmentation has been performed based on Gaussian-based segmentation for automatic segmentation of speech at syllable-level which is automated labelling. Syllable is taken as the basic unit in building Gujarati TTS synthesizer as Indian languages are syllabic in nature. Therefore, in this work, an attempt has been made to reduce these efforts by automatically generating labelled corpus at syllable-level. To that effect, a Gaussian-based segmentation method has been proposed for automatic segmentation of speech at syllable-level. Percentage of Correctness has been taken as measured parameter [15].

As per our knowledge we have cover and consider of mention papers only. It might happen that other researchers may have done work in this field.

## V. CONCLUSIONS

With the above study we can conclude that more development is needed in not only Gujarati but all the regional languages of India. In Concatenative synthesis the task of labeling is most time consuming and tedious process it requires large manual efforts. Furthermore TTS Gujarati synthesiser can be built based on Articulator synthesis and Formant synthesis algorithm by researchers. Accurate Gujarati speech corpus or database and unit selection may improve naturalness and its intelligibility of synthesized speech specifically based on Concatenative synthesis approach.

## REFERENCES

- [1] U R Aida-Zade, C. Ardil and A.M. Sharifova, "The Main Principles of Text-to-Speech Synthesis System", *International Journal of Computer and Information Engineering*, Vol. 7 (3), pp. 395-401, 2013
- [2] Desai Siddhi, Jashin M. Verghese, Desai Bhavik, "Survey on Various Methods of Text to Speech Synthesis", *International Journal of Computer Applications*, Vol. 165 (6), pp. 26-30, May 2017
- [3] Mukta Gahlawat, et al, "Natural Speech Synthesizer for Blind Persons Using Hybrid Approach" *Procedia Computer Science* Volume 41, 2014, Pages 83-88
- [4] Raitio, Tuomo, et al. "HMM-based speech synthesis utilizing glottal inverse filtering." *Audio, Speech, and Language Processing, IEEE Transactions on* vol.19 (1), pp. 153-165, 2011
- [5] Public Document on "Constitution of India 2008" online available at URLs: <https://web.archive.org/web/20140909230437/http://lawmin.nic.in/coi/coiason29july08.pdf>
- [6] Prof. JJ Kothari, CK Kumbharana, "A Phonetic Study for Constructing a Database of Gujarati Characters for Speech Synthesis of Gujarati Text", *International Journal of Computer Applications*, Vol 117(19), pp. 6-11, May 2015.
- [7] J.Sangeetha.et.al, "Text to Speech Synthesis System For Tamil", *International Journal of Emerging Technology and Advanced Engineering*, Volume 3, Special Issue 1, January 2013
- [8] Aiswarya TS,et.al , " Malayalam Text To Speech Synthesis with Emotions", *International Journal of Advance Research in Computer Science and Management Studies* Volume 3, Issue 5, May 2015
- [9] Kaveri Kamble , Ramesh Kagalkar, " A Review: Translation of Text to Speech Conversion for Hindi Language" *International Journal of Science and Research (IJSR)*, Volume 3 Issue 11, November 2014
- [10] Archana Balyan, S.S. Agrwal and Amita Dev, "Speech Synthesis: Review", *IJERT*, ISSN 2278-0181 Vol. 2 (2013) p. 57 – 75
- [11] Mark Tatham and Katherine Morton, *Developments in Speech Synthesis*, John Wiley & Sons, Ltd. ISBN: 0-470-85538-X, 2005).
- [12] Rubeena A. Khan , J. S. Chitode, "Concatenative Speech Synthesis: A Review", *International Journal of Computer Applications*, ISSN 0975 – 8887, Vol. 136(3), pp 1-4, Feb 2016.
- [13] Prof. JJ Kothari, CK Kumbharana, "A Phonetic Study for Constructing a Database of Gujarati Characters for Speech Synthesis of Gujarati Text", *International Journal of Computer Applications*, Vol 117(19), May 2015, pp. 6-11
- [14] Jashin Marina Varghese, Smita Hande, "Text to speech system for Gujarati system for Gujarati language", *International Journal of Advanced Computing and Electronics Technology (IJACET)*, Vol 2(4), 2015, ISSN (Print)-2394-3408
- [15] Swati Talesara, Hemant A. Patil, Tanvina Patel, Hardik Sailor and Nirmesh Shah, "A Novel Gaussian Filter-based Automatic Labeling of Speech Data for TTS System in Gujarati Language", *IEEE proceedings of International Conference on Asian Language Processing*, pp. 139-142, 2013