

Design and Development of Telugu Diphone Database



D.Nagaraju, RJ Ramasree, M.Sri Poojitha, K Nikhitha, M Swetha

Abstract: Diphone database plays a vital role in text to speech system. In this paper, we present the design and development of Telugu diphone database which supports the MBROLA. In the previous article of the same author generation of diphone database for Telugu presents generation of Telugu diphone database [01]. This is the second Telugu diphone database developed by the same author. This database is the upgraded version of first database and few problems are rectified in this version.

Key words: diphone database – Telugu - MBROLA

I INTRODUCTION

Diphone database is the database of sounds. In designing and develop a database of Telugu diphone database there are several steps to follow, first selection of sample words or wav files, which contains diphones, second phone selection, third is diphone selection and extraction of diphone, finally segmenting or equalizing the diphones to make diphone database. In previous article few diphones are not working properly, because their energy levels are not smoothed such few diphones are ui, ai, ea etc. Number of diphones is less in previous diphone database. Above two are major drawbacks in previous database. Theses drawbacks are the motion to make new diphone database in Telugu language.

II RELATED STUDY

A diphone database for Arabic language was designed and developed. Arabic language has 28 consonants, 3 long vowels and 3 short vowels. 102 sentences and 654 words are used to extract 35 phones including pause and 1190 diphones [02]. Romanian language has 1156 possible diphones including 30 allophones [03]. MBROLA supported Sinhala diphone database is designed with 1004 diphones, without considering the allophones [04]. Polish diphone database is created with 1444 diphones and polish extended with 1681 diphones [05]. A Korean diphone database covers 1853 out of the 3977; it covers only 47% of all possible diphones [06].

Lithuanian has 58 phones and two diphone databases are built with possible 5003 diphones [07]. Kurdish language has 37 phonemes, thus it has $37 \times 37 = 1369$ diphones. Diphone database designed with 1300 diphones [08]. Bengali has 44 alphabets and the database designed with 1936 diphones [09].

III TELUGU PHONE LIST AND AUDIO COLLECTION

Telugu language is one of the familiar languages in south India and it is the primary language of the states Andhra Pradesh and Telangana. Telugu people are living across the India and nearly 50 countries in the world. Telugu have 16 vowels and 36 consonants total 52 alphabets, among these there alphabets (ఋ ఋ మఠ్యు అ) are obsolete now [10]. The remaining 49 Telugu phones are considered to build diphone database. Telugu phone set is అ ఆ ఇ ఈ ఉ ఊ ఏ ఏ ఐ ఒ ఓ ఔ అం అః క ఖ గ ఘ జ చ ఛ ఝ ఞ ట ఠ డ ఢ ణ త థ ద ధ న ప ఫ బ భ మ య ర ల వ శ ష స హ ళ క్ష and its transcriptions are a aa i ii u uu e ea ai o oe ou aM a@h ka kha ga gha ~ma ca cha ja Ja ~n Ta Tha Da Dha Na ta tha da dha na pa pha ba bha ma ya ra la va Sa sha sa ha La ksha.

Telugu speech corpus taken from IIT Chennai Telugu mono database wav files [11], these wav files are used to extract diphones. It have 2840 sentence in wav form, among these we used 940 sentences for extraction of diphones.

IV DIPHONE SELECTION AND EXTRACTION

Diphone is an adjacent pair of phones. Diphones are useful in speech synthesis. Pre-recorded diphones are combined to create synthesized speech; the resulting sounds are much more natural than combining just simple phones, because the pronunciation of each phone varies based on surrounding phones. The number of phones in a language is P, and then the number of possible diphones is P². In previous article 106 sentences/622 words are used to extract 1356 diphones and generated the Telugu diphone database. The projected Telugu diphone database is designed with 49 phones and 1806 diphones, these diphones are extracted from 940 sentences. Praat tool is used to extract the diphones from wav files. The General frame work for generation of diphone database is as shown in figure 1.

V EQUALIZATION OF DIPHONES

The energy levels at the beginning and at the end of a segment are modified in order to eliminate amplitude mismatches the energy of all the phones of a given phonemes are set to phones' average value [12].

Revised Manuscript Received on November 30, 2019.

* Correspondence Author

D.Nagaraju*, Associate Professor, Audisankara College of Engineering and technology, Gudur, Nellore, Andhra Pradesh, India -524101

Dr.RJ Ramasree, Professor, Dept of Computer Science, Rashtriya Sanskrit Vidyapeetha, Tirupati, Andhra Pradesh, India

M.Sri Poojitha, Students in Audisankara College of Engineering and Technology, Gudur, Nellore, Andhra Pradesh, India.

K Nikhitha, Students in Audisankara College of Engineering and Technology, Gudur, Nellore, Andhra Pradesh, India.

M Swetha, Students in Audisankara College of Engineering and Technology, Gudur, Nellore, Andhra Pradesh, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Design and Development of Telugu Diphone Database

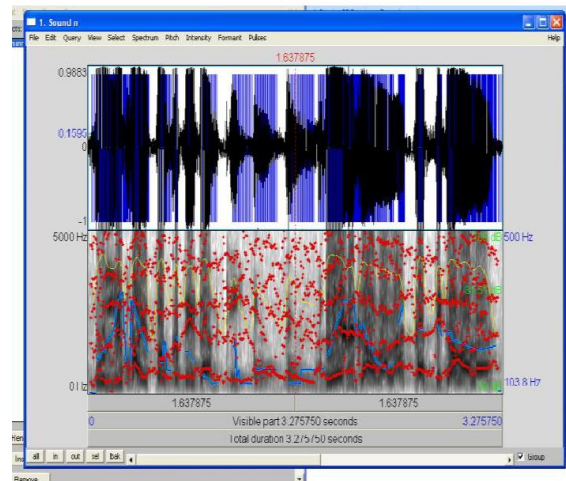
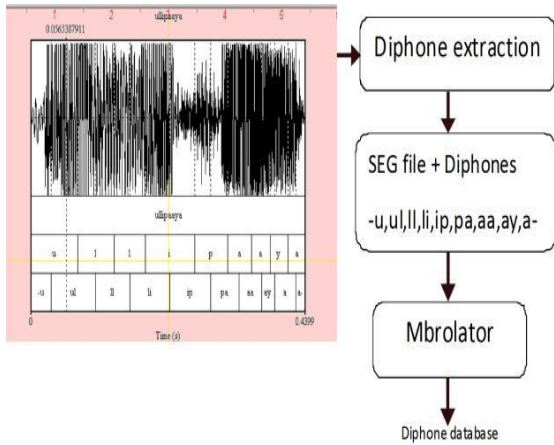
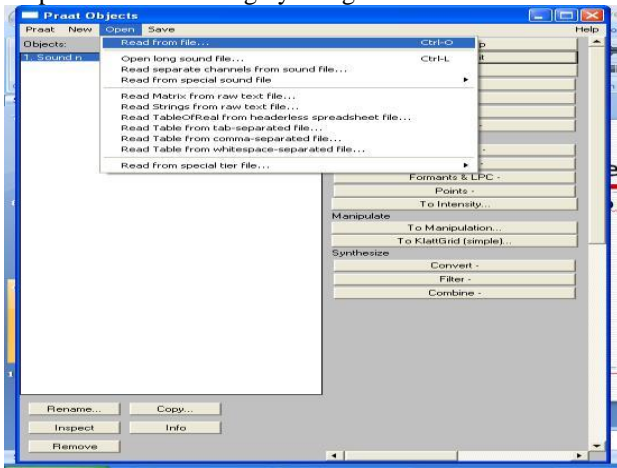


Fig 1. General frame work for generation of diphone database

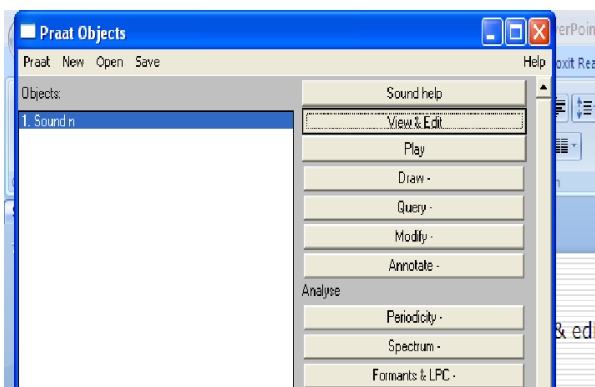
The extracted diphones are then put forward to the evaluation stage in which both database files and the diphone files are evaluated by manually. The annotation diphone files allow manually comparison of the annotations with signals in the diphone files. Mbrulator generates diphone database by taking Database SEG file and all available diphones in the given language.

VI STEP BY STEP PROCEDURE TOWARDS DIPHONE DATABASE GENERATION

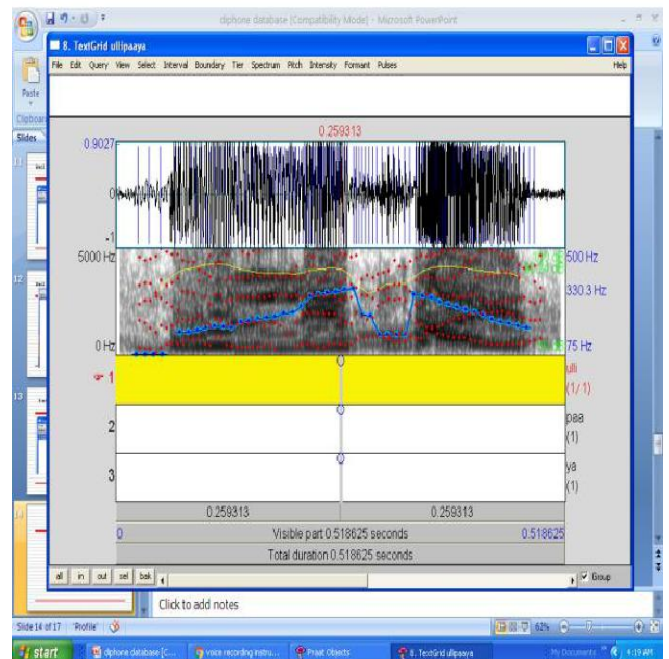
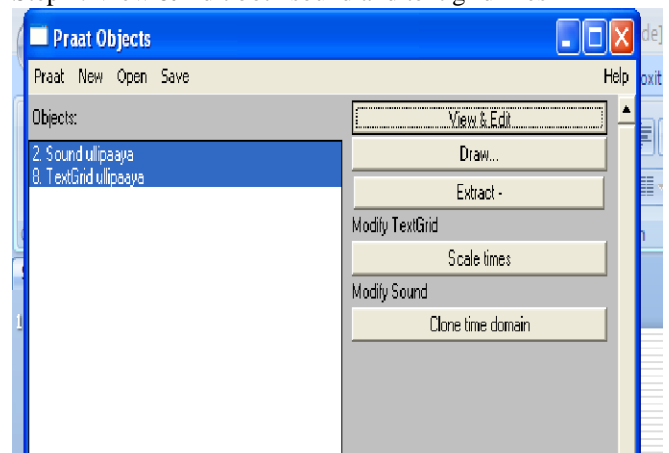
Step1. Wav file reading by using Praat tool



Step2. Click on view and edit command

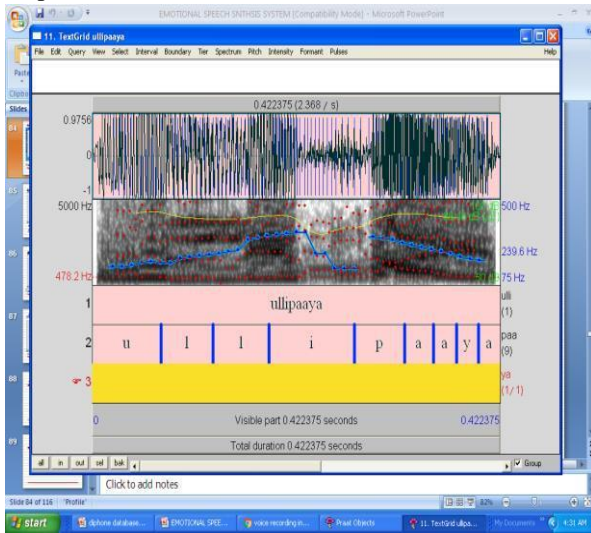


Step 4. View & Edit both sound and text grid files

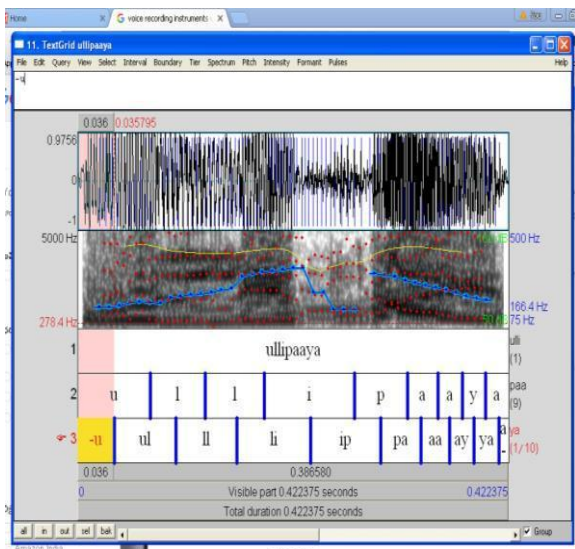


Step3. Create annotation wav file

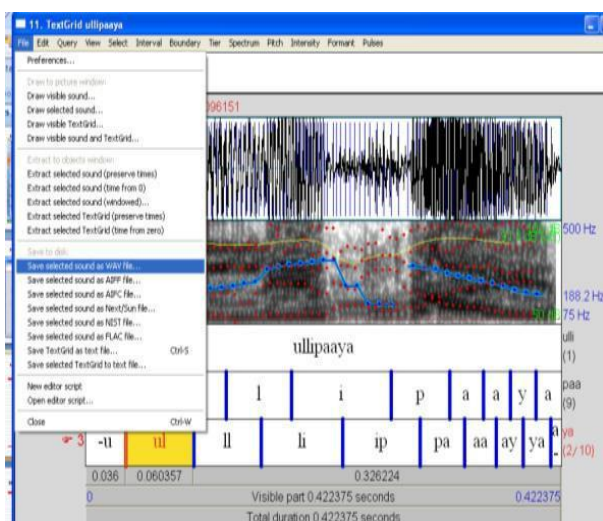
Step 5. Phone selection



Step6. Diphone selection



Step7. Diphone extraction and save selected file as wav



these database it is possible to synthesise a sound of a sufficiently high quality. This can be testified using the examples presented with the data base. Therefore the database alongside the MBROLA software can be used as a module for signal generation in the Telugu speech synthesizer of a high quality.

Previous database have 1356 diphones, it is upgraded up to 1806 diphones and properly not working diphones are upgraded, thus they are working properly.

REFERENCES

1. D.Nagaraju, Dr.R.J.Rama Sree .: Generation of diphone database for Telugu. International Journal of Computational Intelligence Research volume 13, Issue 5, 2015 page no891-898 ISSN:0973-1873research India publications.
2. Aymen El Kadhi, Fadhila Gherri, Hamid Amiri “Building diphone database for Arabic text to speech synthesis system”<http://ieeexplore.ieee.org/document/7233151/>
3. Dragos Burileanu, A. Neagu, C. Burileanu ,Diphone database development for a Romanian language TTS system, Ieee 2002
4. N. G. J. Dias, MBROLA Formatted Diphone Database for Sinhala Language, University of Kelaniya 2015
5. Jolanta Bachan, Efficient_ Diphone_ Database_Creation_for_MBROLA, IEEE 2010
6. Kyuchul Yoon, A prosodic diphone database for Korean text-to-speech synthesis system ,springer 2005
7. Pijus Kasparaitis, Diphone_ Databases _for_Lithuanian_Text-to-Speechsynthesis, inform- atics 2013
8. Wafa Barkhoda Bahram Zahirazami anvar Baharampour Keyvan Shahryari, A comparison between allophone, syllable, and diphone based TTS systems for Kurdish language, Ieee 2010
9. <http://kathak.sourceforge.net/diphone.htm>
10. D.Nagaraju, Dr.R.J.Rama Sree .:Text to Pho algorithm for Telugu. International Journal of Research in Engineering & Advanced Technology, Volume 4, Issue 1, Feb - March, 2016. ISSN: 2320 – 8791 26-28.
11. LDCIL, Linguistic Data consortium for Indian Languages, A linguistic resource repository of all Indian languages for AI/NLP projects. <https://data.ldcil.org/>.
12. D.Nagaraju, Dr.R.J.Rama Sree .:Extraction of diphones for Telugu : Issues and solutions International Journal of Advanced Scientific Technologies ,Engineering and Management Sciences (IJASTEMS- ISSN: 2454-356X) Volume.3,Special Issue.1, March.2017.

VII CONCLUSION

In august 2013 first Telugu diphone database was published. Having properly generated the parameters of the duration and fundamental frequency curves, by means of