ISSUES FACED IN MARYTTS NEW LANGUAGE SUPPORT

Jayanta Kalita
Department of Information Technology
GUIST, Gauhati University, India

Abstract: Development of Text-to-Speech system is still a challenging task for many of the Indian languages. An open source Text-to-Speech system MaryTTS gives platform to create synthesis voice for a new language using various synthesis approaches. Weselect our native language "Assamese"to integrate into MaryTTS using Unit Selection method. We make our compiled Assamese voice freely available at https://drive.google.com/file/d/0B5dIwzIYFl21TVZvVjBxYjc2R1U/view?usp=sharing, which is a necessary step for publishing our voice. As a new user of MaryTTS,we had faced many problems during the process executions. This paper highlights some important issues we faced in the development.

Keywords: MaryTTS, Assamese, New Language Support

1. INTRODUCTION

Text-to-Speech synthesis system is an important requirement for many of the modern applications. The goal of a text to speech system is to convert a given input text into its corresponding speech wave. It is useful in our day to day lives. It helps individuals with incapacities who might be vision weakened. It also useful in eye free environments like while driving, telephone-based voice portals, multi-modal interactive systems etc.Some of the available TTS synthesizers are eSpeak, FreeTTS, MaryTTS, Festival or Flite, PicoTTS etc.

MaryTTS (Modular Architecture for Research on Speech Synthesis) is an open source, multilingual Text-to-Speech Synthesis platform written in Java language. It is a client-server system and can runs on multiple platforms.The latest version of MaryTTS is 5.2 and now it supports German, British and American English, French, Italian, Luxembourgish, Russian, Swedish, Telugu, and Turkish. MaryTTS comes with toolkits for quickly adding support for new languages and for building unit selection and HMM-based synthesis voices [1][2][3].

Marc Schroder and JurgenTrouvain[5] presented the German text-to-speech system MARY, which is a flexible tool for research, development and teaching in the domain of text-to-speech (TTS) synthesis.Though Text to Speech (TTS) is a highly research area in speech synthesis, the Indian languages are still lagging behind. There is no effective Text to speech system available forAssamese language.
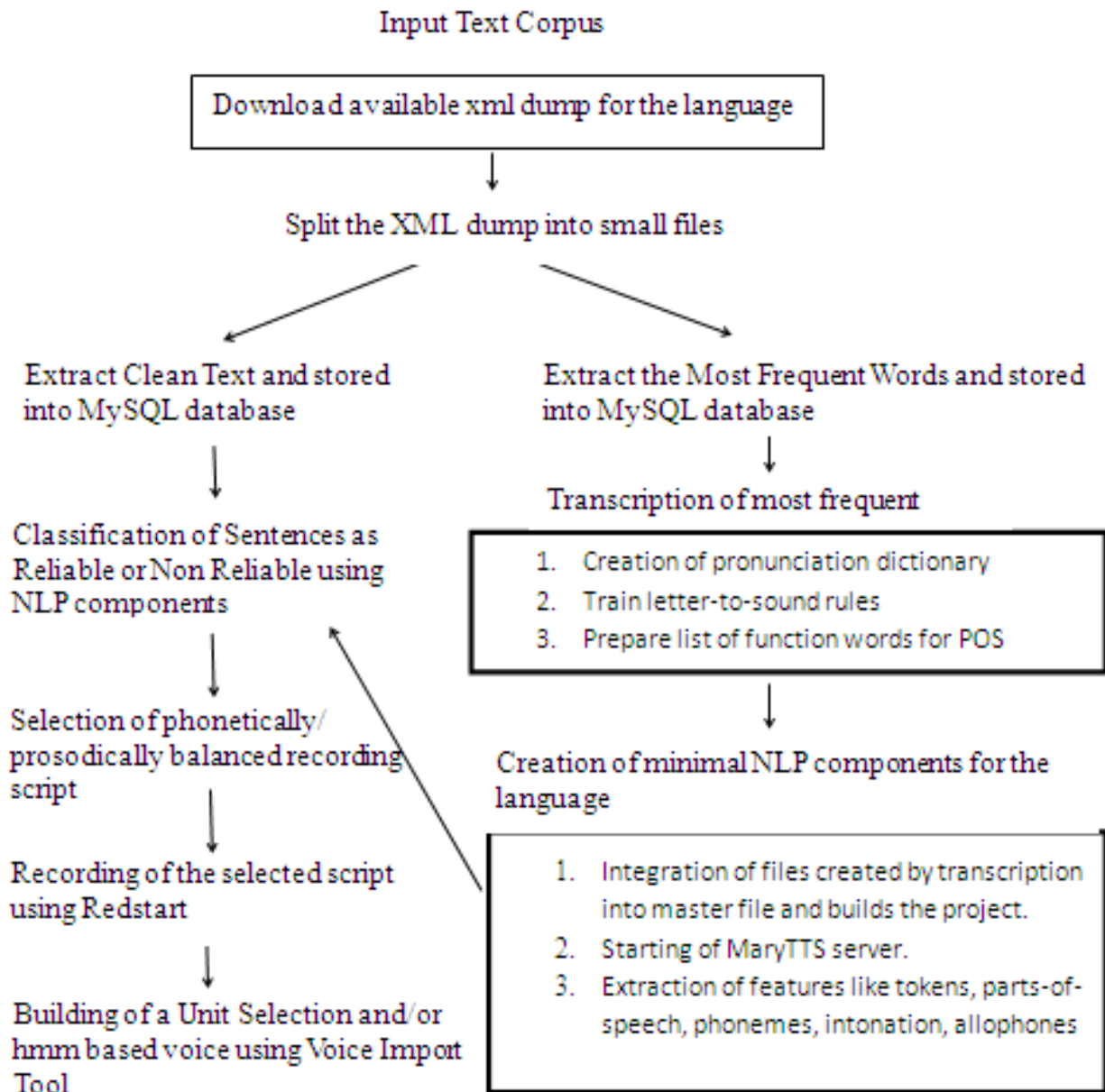
Assamese language is the principal language of the state of"ASSAM" and is regarded as the lingua-franca of thewhole northeast India. The Assamese language is the easternmost member of the Indo-European family. The Assamese language grew out of Sanskrit, the ancient language of the Indiansub-continent. However, its vocabulary, phonology and grammar have substantially beeninfluenced by the original inhabitants of Assam, such as the Bodos and the Kacharis.Assamese uses the Assamese script; a variant of the Eastern Nagari Script, which tracesits descent from the Gupta Script. Developed from Brahmi through Devanagiri, Assamesescript is similar to that of Bengali except the symbols for /r/ and /w/ and highly resemblesthe Devanagiri script of Hindi, Sanskrit and other related Indic languages. The alphabet consists of 12 vowel graphemes and 52 consonant graphemes. Both phonemes and allophones are represented [4].

Here Section 2 shows the workflow of MaryTTS New Language Support and Section 3 highlightssome limitations of MaryTTS New Language Support.

2. WORKFLOW OF MARYTTS NEW LANGUAGE SUPPORT

An open source software developing portal "GitHub" gives platform to add new languages into MaryTTS. The "New Language Support" page shows the steps to add support for a new language.

Input Text Corpus

Download available xml dump for the language

Split the XML dump into small files

Extract Clean Text and stored into MySQL database

Extract the Most Frequent Words and stored into MySQL database

Classification of Sentences as Reliable or Non Reliable using NLP components

Transcription of most frequent

1. Creation of pronunciation dictionary
2. Train letter-to-sound rules
3. Prepare list of function words for POS

Selection of phonetically/ prosodically balanced recording script

Creation of minimal NLP components for the language

Recording of the selected script using Redstart

1. Integration of files created by transcription into master file and builds the project.
2. Starting of MaryTTS server.
3. Extraction of features like tokens, parts-of-speech, phonemes, intonation, allophones

Building of a Unit Selection and/or hmm based voice using Voice Import Tool

Figure 1:Workflow of MaryTTS New Language Support [1]

3. NEW LANGUAGE SUPPORT

MaryTTS has some predefined steps to add support for a new language. But for a new user of Linux it will be difficult to follow the steps due to Linux command line environment and lack of proper guidelines.

The initial step is to setup the environment for MaryTTS in user'slocal machine. It includes compilation of MaryTTS builder tool and creation of wiki data directory. There is a need of some third party software'sin the process of compilation. For example "maven" is use in the voice building process.

In the second step, the user has to download the input text corpus from the Wikipedia link provided. But the link contains database corpus for few languages only. In such cases user has to create the database by itself for the particular language.Again the final output voice of the MaryTTS is highly dependent on the text format of the corpus.

MaryTTS usesMySQLdatabase to store clean text extracted from the corpus.For a user with no experience with MySQL will find very difficult to create a database with all privileges.

After text normalization, MaryTTS transcriptsthe most frequent words. Two files are required in the process. One is an allophone file called "allophone.xy.xml" file and the other is a lexicon file called "xy.txt".Both these files need to be created by the users for their languages with proper format. But there is no properguideline for creation of these two files.

After database selection step, we found that some words and sentences are repeated many times in the database file. Again MaryTTS provides a GUI called SynthesisScriptGUI, which allows a user to delete existing sentences and add new sentences in the database. But for a large database it will be very time consuming task.

In the process of Unit-SelectionVoicebuilding, third party software called HTKLabeler is used for automatic labeling of selected sentences. But HTK Labeler is not further maintained by MaryTTS.So one has to choose other

alternative tool for this process. We used EHMM Labeler in place of HTK Labeler.

4. CONCLUSION

MaryTTS still undergoes continuous development by resolving the issues raised by different users.From our experimental approaches we can conclude that for successful integration of a new language intoMaryTTS one has to have:

(a) Good practice of Linux OS and MySQL.
(b)Select a phonetically and prosodic balanced text corpus.
(c)Select a recent stable version of MaryTTS.
(e) Should have clear and strong knowledge of phonetics used in their languages.

Here we highlight some issues we had faced during integration of our native language "Assamese" intoMaryTTS.Hope this paper will help other researchers and developers who are following the "New language support" to solve similar issues.

5. ACKNOWLEDGEMENT

6. REFERENCES

[1]. https://github.com/MaryTTS/MaryTTS/wiki/New-Language-Support
[2]. http://mary.dfki.de
[3]. https://mary.opendfki.de
[4]. http://www.iitg.ernet.in/rcilts/assamese.html
[5]. Marc Schroder, JurgenTrouvain;" The German Text-to-Speech Synthesis System MARY:A Tool
[6]. for Research, Development and Teaching", International Journal of Speech Technology ,6-4, 365- 377 (2003)