Maithili Text-to-Speech System

Amit Kumar Jha and Piyush Pratap Singh
Center for Information and Language Engineering
Mahatma Gandhi Antarrashtriya Hindi Vishwvidyalaya
Wardha, Maharashtra 442001, India
jhaamit036@gmail.com, piyushpratapsingh@rediffmail.com

Pankaj Dwivedi
Educational Technology Unit
Central Institute of Indian Language
Mysuru, Karnataka 570006, India
pankaj.linguistics@gmail.com

Abstract - This paper discusses development of TTS system for Maithili language. Speech corpus spanning 5 hours borrowed from LDCIL, CIIL Mysore, and also of 3 hours is collected from native speakers in studio environment. As most Indian languages including Maithili are syllabic in nature, concatenative method is used for the purpose of speech generation taking syllable as a basic unit. To enhance naturalness of speech out, 1055 most frequently occurring words have been recorded and stored. The system supports UTF-16 for text input. C#.NET is used for development of interface. The speech database consists of 930 syllable (C*V) in total. Each position has 300 syllables and 10 independent vowels. 930 units of speech data is built from all three positions. Subjective Evaluation, MOS and MRT, are conducted by 10 native speakers. The quality of synthesized speech in terms of intelligibility and naturalness is evaluated to be approximately 84 percent. The relevance of the work lies in the fact that no TTS system exists for Maithili Language till date.

Index Terms - Text-to-Speech System, Speech Processing, Maithili, Concatenate method.

I. INTRODUCTION

A Text-to-Speech System (TTS) converts a raw text into human speech sounds. It can be of significant aid to communication for visually impaired people other than its role in telecommunication, industrial and educational applications. Government of India initiated development of TTS systems for Indian languages through TTS consortium project under MeitY [1]. 13 Indian languages, namely Hindi, Gujarati, Telugu, Marathi, Tamil, Odia, Malayalam, Bengali Assamese, Kannada, Bodo, Manipuri, and Rajasthani were covered under this TTS consortium for development of TTS system [2]. Since then reasonable advancement has been made in form of minor and major research & development initiatives in TTS technologies for these and some other languages. However, many Indian languages, listed in 8th schedule of the constitution, having more number of speakers and wider geographical distribution are still devoid of any such research as far as TTS technologies are concerned. Maithili (ISO 639-3 - mai) is one of such Indian languages.

Since Indian language characters are complex in nature, there is no straightforward approach to build TTS system for Indian languages as compared to English or other European languages [3]. TTS systems, for various Indian languages, have been developed using various approaches such as articulatory synthesis [3], formant synthesis [4], unit selection synthesis (USS) [5], [6] and HMM based speech synthesis (HTS) [7]. In articulatory synthesis, human vocal organs are modeled using anatomic or physiologic parameters; while in formant synthesis predefined acoustic properties such as fundamental frequency, voicing, and noise level are processed [8]. USS based TTS systems are more successful in generating speech closer to natural speech. USS based concatenative method in Festival framework has been vastly used for different Indian languages including Hindi [9], [10], Telugu [9], [11], Bengali [12] [13]. However, one of the main drawbacks of concatenate speech synthesis is its limitation to a single voice [14]. HTS system gives smaller footprint compared to USS system due to its statistical modeling. The statistical representation also enables transforming voice characteristics, speaking styles, and emotions, by mimicking voices, mixing and controlling voices [7]. In Festival framework TTS system, phone or diphone is used as a basic unit. Deep Neural Network (DNN) based method eliminates the need for a grapheme to phoneme conversion; rather, it synthesizes speech directly from the UTF-8 based script [15]. Fig 1 depicts milestone year(s) in the development and advancement of TTS technologies under various approaches. The leftmost side lists introduction of mechanical TTS system by Kratzenstein in 1779, and rightmost side shows use of DNN in development of TTS systems.

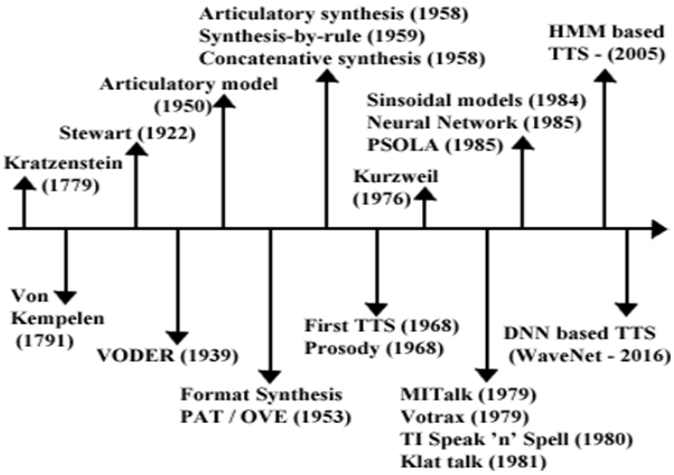

Fig.1. Milestones in the field of TTS

978-1-7281-2472-8/19/$31.00 ©2019 IEEE

There are several institutes and research organizations engaged in the field of TTS system development for Indian languages. Some of the main works are listed as below: -

- IIT Mumbai - Vani Framework
- C-DAC Kolkata - Bengali TTS
- HP Labs - Hindi TTS
- Simputer Trust - Dhvani TTS
- IIT Kharagpur - Shruti TTS
- C-DAC Bangalore - Matrubhasha
- IIIT Hyderabad - Telugu TTS
- C-DAC Thiruvananthapuram - Malayalam TTS (SUBHASHINI)
- IISC Bangalore - Thirukkural & Vaachaka
- TIFR, Mumbai - Indian English TTS

This paper is divided into five sections. Section I introduces the previous works and reviews literature in the concerned area. Section II focuses on database creation and sound system of Maithili language within scope of the present work. Section III explains proposed algorithm, while IV provides detailed discussion of results. Section V briefs about challenges, limitations and future scope of the present work.

II. MAITHILI LANGUAGE AND DATABASE CREATION

Maithili (EGIDS 0-4) is an Indo-Aryan language. It is primarily spoken in the state of Bihar, and also in some other parts of India and Eastern Nepal approximately by more than 30 million people in total [16]. It is primarily written in Devanagari script. Kaithi or Mithilakshar (also known as Tirhuta) script may be used in highly limited domains. Newari script is used to write Maithili by the newar tribe of Nepal. Kaithi script is used by the Kayasthas or bookkeepers in writing Maithili language. After being incorporated in 8th schedule in the year 2002, it slowly is picking up to be used in government, education and other official communicative contexts.

There are total 16 phonologically distinctive vowel segments in Maithili. Of these 16 vowels, 8 are oral vowels [i. e. æ a ɔ o u] and 8 are corresponding nasal vowels [i` ẽ æ̃ ã ɔ̃ õ ũ] [17]. Two distinctive oral diphthongs /əi/ and /əu/ are also found. Although length of vowels is not phonemic, Maithili does use both long and short allophonic variants of all eight vowels. The long allophonic variants occur primarily in stressed syllables; the short ones occur in unstressed syllables [17]. A total of 33 consonantal segments [p, p^h, b, b^h, t̪, $t̪^h$, d̪, $d̪^h$, ʈ, $ʈ^h$, ɖ, $ɖ^h$, k, k^h, g, g^h, ʧ, $ʧ^h$, ʤ, $ʤ^h$, s, h, m, n, ŋ, ʃ, ɳ, ʂ,w (v), j, r, ɽ, l,] are used in written form, only 30 segments are realized in speech of a native speaker, however. The consonants [ʃ, ɳ, ʂ] are replaced by [s, n and s], respectively in speech. For example, <ʃaːm> → [saːm] 'evening', <baːɳ> → [baːn] 'arrow', and <kəʂt> → [kəst] 'pain'. The sound [ɽ] is also realized as [r] is intervocalic and syllabic boundary positions. For Example, <$gʰoɽaː$> → [$gʰoraː$] 'horse' and < kəɽ.ək > →[kər.ək] 'strict'.

Maithili is not a phonetically perfect language yet its consonants and vowels mostly have one-to-one correspondence between written and spoken form. Like Hindi, there are some issue such as schwa (ə) deletion & context dependent pronunciation of schwa sound, syllabic boundary and pronunciation of anusvara and chandra-bindu in Maithili phonology that needs to be duly addressed. Maithili exhibits four types of syllabic structure V, CV, VC and CVC. Its phonotactics allows syllabic margins to accommodate a maximum of two consonants. Syllables having more than two consonants are broken down by vowel insertion with few exceptions as attested in many of its sister linguistic varieties.

Database preparation is an essential and important component in any TTS system which is built on concatenative method. However, this preparation may vary with one language to another, and also on availability of language resources. Preparation of database starts with collection of variety of phonetically balanced text data (corpus). These texts are speech recorded by speaker(s) having native proficiency of language in highly controlled recording studio environment. To prepare a rich database for our TTS system, text data from different domains like children stories, literature, science, tourism, politics, history, daily affairs, drama, poetry, etc. were adequately covered. The main sources of the data were published books, newspapers, local periodicals & magazines, and web pages & blogs. Among these sources, books published in Maithili by NCERT, Bihar, Patna of Maithili language for class 1-8 were very helpful. Syllable combinations which could not be covered from these sources were taken from Maithili Dictionaries. Kalyani shabdkosh proved to be extremely useful among others [18]. To cover old and folk terminologies, about 120 oral and folk narratives (stories and legends) were first text recorded, and thereafter read by native speaker(s).

Recording is done at sampling frequency of 16 KHz/16 bits. PRAAT software is used in truncation and annotation of the sounds files at sentence, word, and syllable levels. Syllables have been stored covering all three different possible positions i.e. initial, middle and final to account for maximally possible phonetic coverage. For example, an syllable 'ता' [] might be cut from taken from words 'ताब' [t̪ab] <then>, 'लताम' [lətam] <guava>, and 'सरोता' [sərɔt̪a] <nutcracker> covering initial, middle and final positions. Our speech database consist of 930 syllable (C*V). Each position has 300 syllables and 10 independent vowels. Hence, from all three positions of [(300*3) + (10*3)] = 930 units of speech data has built.

III. PROPOSED ALGORITHM

To develop this TTS system, concatenate Unit Selection Synthesis (USS) technique is used. In terms of naturalness, USS generates better speech as it uses small amount of Digital

Signal Processing (DSP) to speech recorded data. Often DSP makes speech less natural; to smoothen the waveform, some systems anyway apply small amount of signal processing at the point of concatenation. Since Indian languages including Maithili are syllable-centered [19], syllables & polysyllables, rather than phonemes, have been taken as basic unit of concatenation in our work. Considering phoneme as a basic unit of concatenation may result in low quality speech output due to large concatenation points. In database, syllables are stored in all three initial position, middle position and final position.

Fig 2 describes workflow chart of our TTS system.

1) Input Maithili text written in Devanagari script using UTF-16.
2) Inputted text is normalized with the help of three algorithmic modules written in C# and SQL.
3) Inputted text is segmented into sentence level. Afterwards, it is segmented into word level using white space.
4) A word level search is done in database and if it is found then corresponding speech file is added into playlist. Else, the word is broken into corresponding syllables and corresponding syllables files are searched and added in playlist.
5) Found speech units are concatenated in playlist using digital signal processing.
6) Add prosodic features to the speak file according to the types of sentence.
7) Play the sound of playlist.

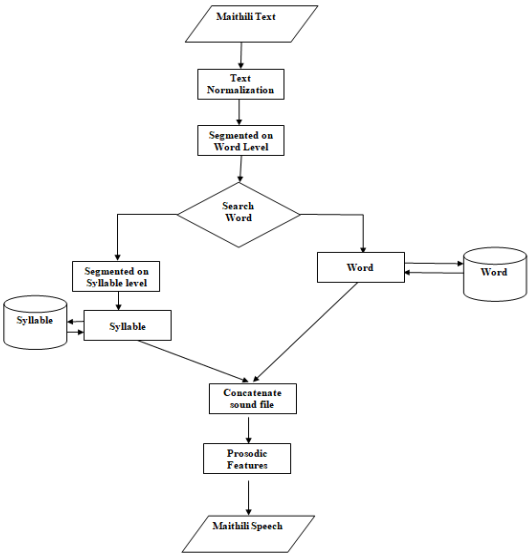

Fig.2. Workflow of TTS System for Maithili Language

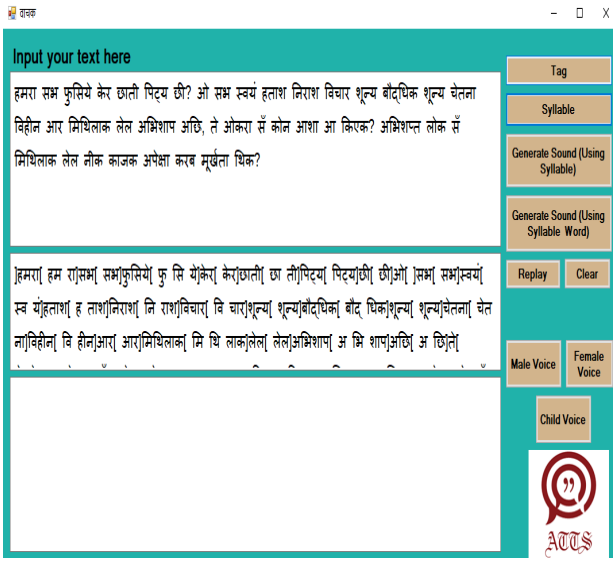

Fig.3. Screenshot of TTS for Maithili

A. *Text Preprocessing*

Building of any TTS system requires dealing with unrestricted or real text, which in turn requires dealing with Non-standard Words (NSWs) such as numbers (time, ordinal, cardinal, floating point, year, phone/mobile number), abbreviation, dates, URLs, acronyms, etc. These NSW must be normalized, i.e., expanded for their correct pronunciation given context. The steps of normalization are as below –

1) Tokenize the input text into words based on white space and special symbol such as, purn viram (full stop), semicolon, comma, colon, question mark, exclamation mark, etc.

2) Identify the NSW tokens such as abbreviation, acronyms, number, fractions, ratios, symbols, dates, time, etc.

3) Classify the tokens as abbreviation, acronyms, numbers, symbol, date, URL, etc.

4) Convert Non-standard words to standards words by corresponding expansion rules and developed lexicon.

Normalization of numbers requires more attention than other letter strings. Numbers are pronounced as individual digits or as a whole, depending on the context. Expansion of a number demands making of a decision for the number being cardinal, ordinal, date, time, currency or referring to anything else. For example, if a given number is *currency* such as rupee, dollar or pound, the input number is treated as a sting. A distinct module is written for the expansion of the *ordinary number*. The control function chooses different conditions according to the length (number) of number string. If the number of digit is 3 then it must be in hundreds (for example 546 – पाँच सौ

छियालीस). If a number string is 4 or 5 places long, it must be in thousands or ten thousands, respectively. In such case, the above module is called twice. First, number of thousands is converted into words (1 or 2 digits) and second remaining three digits are converted. For example, if the number is 13656, in the first call number 13 is processed, and in the next call 656 is processed. Similarly, if number string is 6 or 7 places long, it should me in lakh or ten lakhs then the module is called thrice. If the given number string is not an ordinary number, a decision is taken to determine the type of number. In *decimal number* (32.5478), the number before the dot is pronounced as one number, while those after the dot are pronounced as isolated digit. Number string indicating *dates* are identified with the help of delimiters '/' or '-', or '.' (Ex: 13-02-2019, 11/12/1989, or 25.09.2001). All the three values (date, month, and year) are extracted from the input string and processed separately. For example - 11/03/2007 will normalized into ग्यारह मार्च दु हजार सात. The MM is normalized as the name of the month according to the order of month. For example 01 is normalized as जनवरी, 02 is normalized as फरवरी, 03 is normalized as मार्च, etc. Number string indicating *time* is identified with colon (:), which is normalized to pronounced as [बजकर]. Digits both before and after colon are pronounced as single numbers with addition of word [मिनट] after the number spoken after colon. The number string absent any of the above parameters, are normalized to pronounce as individual digits in isolation. Therefore, (*mobile/ISBN/Barcode*, etc.) no. 9612543687 is normalized into corresponding spoken form of their individual digits, and therefore pronounced as "नौ छह एक दु पाँच चारि तीन छह आठ सात".

Table1. Normalization of Number Strings

Input NSWs	Normalized Output
138485	एक तीन आठ चारि आठ पाँच
48.36	अरतालीस दसमलब तीन छह
12-03-2010	बारह मार्च दु हजार दस
12-03-1996	बारह मार्च उन्नीस सौ छियानवे
16.06.1857	सोलह जून अठारह सौ संतावन
11:40	ग्यारह बजकर चालीस मिनट
92583	नौ दु पाँच आठ तीन

Like Hindi and other Indian languages, Maithili text offers peculiarities such as interspersed English word using roman script, number convention (Arabic, Roman and Devanagari), foreign language tokens. (e.g. AK-47, MIG-29). This may be achieved by replacing raw character string by a normalized letter string which corresponds to the respective pronunciation of the acronym or the abbreviation, or special symbol or characters. For example - Dr is normalized as doctor; M.A. as Master of Arts; % as 'Pratisat', @ as 'at the rate' etc. in Maithili.

Table 2. Normalized Pronunciation of Abbreviations and Acronyms

NSWs	Normalized form	Gloss
प्रो.	प्रोफेसर	Professor
भाजपा	भारतीय जनता पार्टी	BJP
AK-47	एके सैतालीस	AK-47
%	प्रतिसत	Percentage
$	डाउलर	Dollar
65 Kg	पैसठ किलोग्राम	Sixty five Kilogramme
20 L	बीस लीटर	Twenty Litters

To develop lexicon(s) for frequent NSWs unique to Maithili text (morphology), Structure Query Language (SQL) is used. A table consisting of two fields, called non standards words and its equivalent standards words. Algorithm identifies the NSW token and searches it in the database, i.e., lexicon of NSWs, and then returns normalized corresponding spoken form of the token.

In Maithili, a limited set of tokens, mostly auxiliaries and Nouns, are pronounced differently than their corresponding written form purely due to socio historical reasons such as adoption for Devanagari Script over Kaithi for writing Maithili. A separate a lexicon is developed for these words. Like NSWs, algorithm identifies these tokens and searches them in database, then returns normalized corresponding spoken form of the token. Table 3 lists some of such words having phonetic discrepancy.

Table 3. Normalized Pronunciation of some NSWs

NSW	Normalized form	Gloss
अछि	अइछ	is 'Hon'
छथि	छइथ	was ' Hon'
छलन्हि	छलइन्ह	were 'Hon'
जाथि	जाइथ	go-PST
राति	राइत	night
लोकनि	लोकइन	people

To increase the naturalness of the TTS system, 1055 most frequently occurring words have been recorded and stored in separate lexicon.

IV. RESULT AND DISCUSSION

Evaluation of TTS system depends on the purpose of the evaluation, and also on the specific purpose for which the synthetic generated voice is intended. Three key approaches used for the purpose of evaluation are: 1 User testing, 2 Objective measures, and 3 Feature comparisons. While these three approaches may seem different from one another, they complement one another in actuality, however. Different evaluation approach may be carried out at a different stage in

the development and testing of a TTS system. Our Maithili TTS system is evaluated based on subjective user measurement/testing approach to see the naturalness and intelligibility of the output of the Maithili TTS. Mean Opinion Score (MOS) from 10 users was calculated on test data. The test data covered normally found variations in language such as numbers, abbreviations & acronyms, symbols, English words written in Devanagari script, etc. from different domains namely news, stories, sports etc.

The 10 native (5 Male and 5 Female) speakers having native Maithili language are selected for the evaluation of our TTS system. Details regarding their age, gender, educational qualification, and exposure to other language are given in Table 4.

Table 4. Evaluators' details

Name	Age	Gender	Qualification	Other language
E1	32	Female	Matric	Hindi
E2	26	Male	Graduate	Hindi
E3	21	Female	Graduate	Angika
E4	23	Female	Graduate	Hindi
E5	34	Male	Master	Hindi
E6	24	Male	Graduate	Hindi
E7	24	Male	Master	Angika
E8	26	Female	Graduate	Hindi
E9	33	Female	Master	Angika
E10	25	Female	Graduate	Hindi

It is ensured that none of evaluators have history of hearing loss or impairment. Evaluators also went through a small training session to make themselves familiar with the evaluation metrics and working of TTS. The evaluation process has been conducted in noise free studio environment (NR 10 with maximum variation up to NR 15*)* using high quality headphone (Brand *Audio-Technical* model *ATH-M50X*). 10 sample sentences covering different domain are as following:-

1) अमावाश्याक अन्हरिया राति मे ट्रेन खटर खटर के आवाज करैत तीव्र वेग सँ जा रहल छल।

2) दुनियाँ दोसर भाग मे आम'क महत्व शायदे एत्तेक भऽ सकैत अछि।

3) अहाँ हमरा बुरबक बुझैत छी ?

4) बिहार सरकारक कला संस्कृति एवं युवा विभाग 18 अक्टूबर 2016 सँ लSक 17 अक्टूबर 2017 धरि बिहार कला दिवसक रुपे मनेबाक घोषणा कएलाह।

5) लगभग 4 साल होमय जा रहल अछि।

6) बढैत गेल बीमारी, घटैत गेल इलाज।

7) ओना तऽ ट्रेन 2:30 में खुलतै मुदा भीड़ अभीस बेसी छल।

8) उ ओते बात करे छै जते भात मिलै छै।

9) तखन ओ कियैक 21 लाख पुरस्कार राशि बंटताह ?

10) पटना मे शुरू भेल पहिल विश्व महिला कबड्डी प्रतियोगिता।

MOS technique is used to assess the quality of the speech score from 1 to 5, where 5 refers to Excellent, 4 stands for Good, 3 is for Fair, 2 is for Poor, and 1 refers to Bad on the scale of linguistic nativity of Maithili.

Based on the MOS table 5, it is clear that the quality of the TTS system for Maithili language is 4.2, i.e. 84 percent, which is a good for the TTS system of any Indian languages.

Table 5. MOS table for Quality analysis of TTS system for Maithili

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	Avg.
S1	5	3	4	3	4	4	5	4	3	5	4
S2	4	3	3	4	5	5	3	4	5	4	4
S3	3	4	5	4	4	5	4	5	4	5	4.3
S4	5	3	4	3	3	4	3	4	3	4	3.6
S5	4	5	4	4	3	3	4	5	3	5	4
S6	4	4	5	4	5	4	5	3	5	4	4.3
S7	5	4	4	3	5	4	5	4	4	4	4.2
S8	4	5	4	4	3	5	3	5	3	3	3.9
S9	5	4	5	4	4	4	3	4	4	4	4.1
S10	3	4	5	3	4	5	4	3	3	5	3.9

Chart 1 represents that S4 has the lowest acceptability among the evaluators, while S3 and S6 have the most acceptability percentage. It's fair to conclude that naturalness of the any synthetic voice also depends on length of voice chunk, presence of foreign linguistic tokens, etc. The S4 has maximum linguistic chunks, borrowed Hindi words, also two number strings joined by conjunction. S3 has lowest linguistic chunks and all words are native to Maithili. S4 has second lowest linguistic chunks referring to a popular quotation native to Maithili.

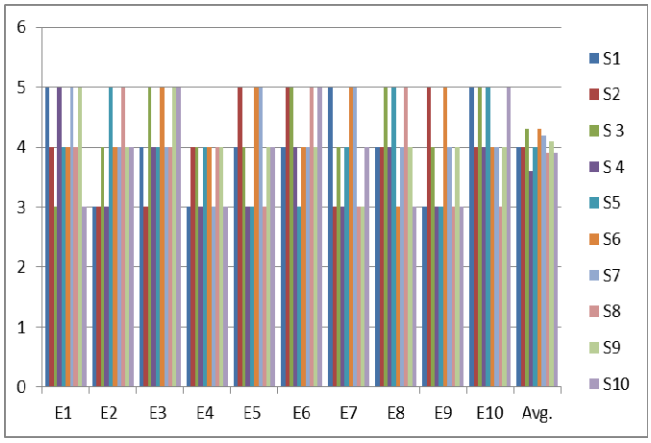

Chart 1. MOS Chart for Quality analysis of TTS system for Maithili

After the subjective evaluation (MOS), segmental evaluation, based on Modified Rhyme Test (MRT) is also conducted. MRT judges whether the evaluators can differentiate the rhyming words or number in a given language. Overall results of MRT showed that 7 out 10 native Maithili speakers were

able to instantly differentiate the rhyming words based on near minimal features such as -voiced/+voiced and +nasal/-nasal on more than 82% occasions.

V. CONCLUSION

The output of this system is satisfactorily close (84%) to the natural Maithili speech. Quality of synthesized speech is acceptable and intelligible.

Some of the main challenges that we faced in building present TTS system are that Maithili is under-resourced language. When the target language is under-resourced or low resourced, it is likely to be influenced by one or more embedded foreign languages [20]. In our database Hindi and English (both single tokens and strings) are unavoidably and frequently mixed. Also, noticeable prosodic differences in some speech samples were found due to intra-dialectal and inter-dialectal differences. These prosodic differences are more frequent in samples have more instances of code mixed and code switched speech, which in turn compromise the overall naturalness and quality of synthesized speech. The present TTS system can be made more robust with implementation of prosodic features.

REFERENCES

[1] TDIL TTS, "Indian Language Technology Proliferation and Deployment Centre: Text-to-speech," available: http://tdil-dc.in/.

[2] H. A. Patil, et al, "A syllable-based framework for unit selection synthesis in 13 Indian languages," in 2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (OCOCOSDA/ CASLRE), Nov 2013, pp. 1–8.

[3] J. Benesty, M. M. Sondhi, and Y. A. Huang, Springer Handbook of Speech Processing. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2007

[4] W. Lawrence, "The synthesis of speech from signals which have a low information rate," Communication Theory, pp. 460–469, 1953.

[5] D. H. Klatt, "Review of text to speech conversion for english," The Journal of the Acoustical Society of America, vol. 82, no. 3, pp. 737–793, 1987. [Online]. Available: http://scitation.aip.org/content/asa/journal/jasa/82/3/10.1121/1.395275

[6] A. Hunt and A. Black, "Unit selection in a concatenative speech synthesis system using a large speech database," in Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on, vol. 1, May 1996, pp. 373–376 vol. 1.

[7] H. Zen, K. Tokuda, and A. W. Black, "Statistical parametric speech synthesis," Speech Communication, vol. 51(11), pp. 1039–1064, 2009

[8] Aparna N S and Shreekanth T, "Text to Speech Synthesis of Hindi Language using Polysyllable Units", National Conference on Power Systems & Industrial Automation (NCPSIA 2015).

[9] S.P. Kishore, R. Sangal and M. Srinivas, "Building Hindi and Telugu Voices using Festvox", Proceedings of the International Conference On Natutal Language Processing 2002 (ICON-2002), Mumbai, India, 2002.

[10] A.G. Ramakishnan, K. Bali, P.P. Talukdar and N.S. Krishna, "Tools for the Development of a Hindi Speech Synthesis System", In 5th ISCA Speech Synthesis Workshop, Pittsburgh, 2004, pp. 109-114.

[11] C. Kamisetty and S.M. Adapa, Telugu Festival Text-to-Speech System. Retrieved from: http://festivalte.sourceforge.net/wiki/Main_Page, 2006.

[12] Alam, F., Nath, P. K., & Khan, M. (2007). Text to speech for Bangla language using festival.

[13] Das Mandal, S. K., & Pal, B. (2002). Bengali text to speech synthesis system a novel approach for crossing literacy barrier. CSIYITPA (E)

[14] https://www.explainthatstuff.com/how-speech-synthesis-works.html 17:05, 26-03-2019

[15] Sharma, B., Adiga, N., & Prasanna, S. M. (2015, November). Development of Assamese text-to-speech synthesis system. In TENCON 2015-2015 IEEE Region 10 Conference (pp. 1-6). IEEE

[16] https://www.ethnologue.com/language/mai, 16:52, 17-02-19

[17] Sunil Kumar Jha, "Maithili Some Aspect of Its Phonetics and Phonology", Motilal Banarsidass publishers private limited, Delhi, 2001

[18] Govind Jha, Eds., Kalyani Kosh "A Maithili-English Dictionary, Maharajadhiraj Kameshwar singh Kalyani foundation", Darbhanga, 1999

[19] R. Venugopalakrishna Y, et.el, Design and Development of a Text-To-Speech Synthesizer for Indian Languages

[20] P. Dwivedi and S. Kar. "On documenting low resourced Indian languages insights from Kanauji speech corpus." Dialectologia: revista electronica. pp 67-91, 2017. Available: https://www.raco.cat/index.php/Dialectologia/article/viewFile/328489/419096