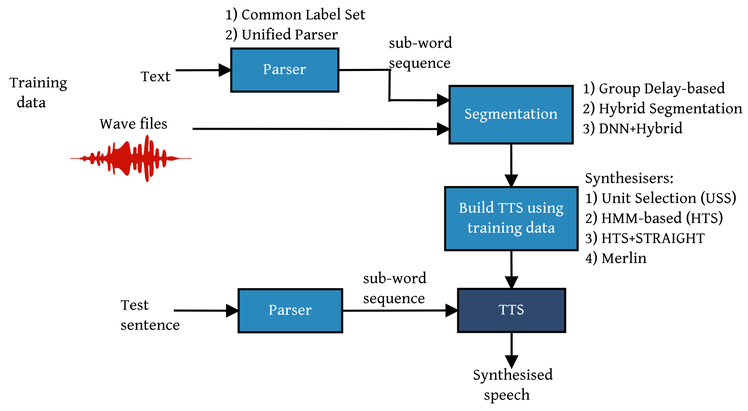
Speech Synthesis in Indian Languages and Future Perspectives

Arun Baby, Anju Thomas, Jeena J Prakash, Anusha Prakash, Hema A. Murthy
arunbaby@cse.iitm.ac.in
Indian Institute of Technology, Madras, India

Introduction



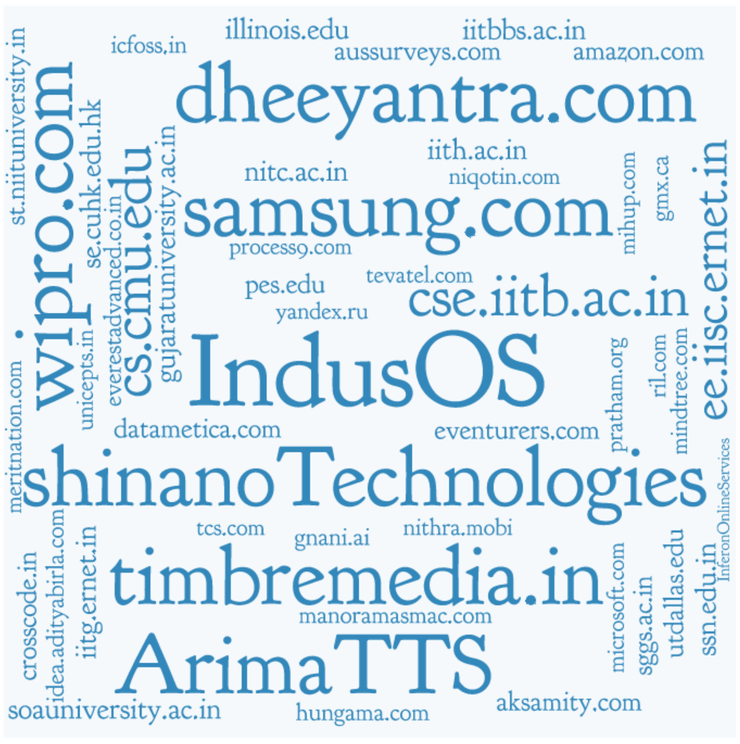
- Text to speech synthesis system converts given input text to speech utterance.
- Speech corpus for 13 Indian languages, each containing 40 hours of data(native and English of male and female speakers respectively).
- Unit selection/statistical parametric speech synthesis systems for Indian languages.

To learn more, visit our website IndicTTS @ *https://www.iitm.ac.in/donlab/tts/*

Phase I

- Syllable based unit selection synthesis systems for 6 Indian languages.
- Integration of TTS with ORCA/NVDA.
- Training workshops for visually challenged persons at different places.

Beneficiaries



Future Perspectives

- Prosody based on context.
- Dynamic IVRs.
- Health care, weather reports.
- Google maps enabled with Indic voices.
- Android based applications to aid persons with cerebral palsy, visual impairments.

Phase II

	Label	IPA	Hindi	Marathi	Bengali P	Bengali G	Tamil	Malayalam	Telugu
Vowels	a	a	अ	अ	–	–	அ	അ	అ
	ax	ɔ	–	ऑ	অ	অ	–	–	–
	aa	a:	आ	आ	আ	আ	ஆ	ആ	ఆ
	i	ɪ, i	इ	इ	ই,ঈ	ই	இ	ഇ	ఇ
	ii	i:	ई	ई	–	ঈ	ஈ	ഈ	ఈ
	rq	-	ऋ,ॠ	ऋ,ॠ	ঋ,ৠ	ঋ,ৠ	–	ഋ	ఋ,ౠ
Stop Consonants	c	ʧ	च	च	চ	চ	ச	ച	చ
	ch	ʧʰ	छ	छ	ছ	ছ	–	ഛ	ఛ
	cx	ʄʃ	–	च	–	–	–	–	–
	jx	ʄʒ	–	ज	–	–	–	–	–
Semi-Vowels	y	j	य,य़	य,य़	য়	য	ய	യ	య
	l	l	ल	ल	ল	ল	ல	ല	ల
	lx	ɭ	–	ळ,ऴ	–	–	ள	ള	ళ
	w	ʋ	व	व	–	ওয়	வ	വ	వ
Fricatives	sh	ʃ	श	श	শ,য়	শ	–	ശ	శ
	s	s	स	स	স	স	ஸ	സ	స
	khq	x	ख	ख	–	–	–	–	–
	z	z	ज़	ज़	জ়	–	–	–	–
	f	f	फ़	फ़	–	ফ়	ஃப	–	–

Fig 1: Common Label Set

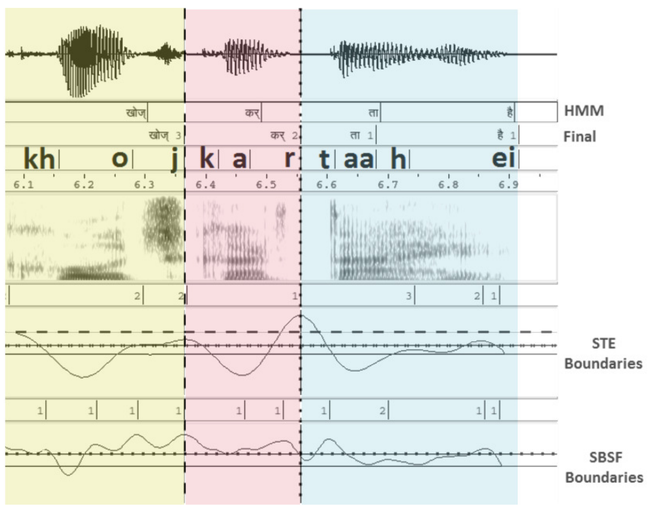


Fig 3: Segmentation

Screenshot 1: IndicTTS Website

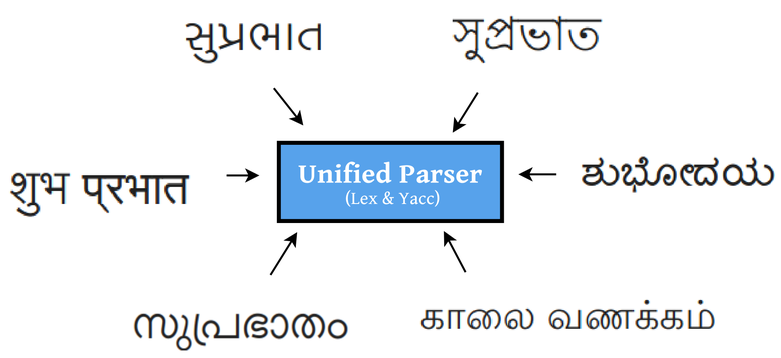


Fig 2: Unified Parser

Fig 4: Android Applications

Download Type	Count
Database	2519
CLS	403
Unified Parser	381
Hybrid Segmentation	263
Pruning	149
Voices	803
Android Applications	576
Synthesis Documents	1470

Screenshot 2: Download Statistics

References

- Anila Susan Kurian et.al, "Indian Language Screen Readers and Syllable Based Festival Text-to-Speech Synthesis System", Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies, Association for Computational Linguistics, pp. 63 - 72,July, 2011, Edinburgh, Scotland, UK.
- Arun Baby, Nishanthi N L, Anju Leela Thomas, and Hema A Murthy, "A unified parser for developing Indian language text to speech synthesizers", 19th International Conference on Text, Speech and Dialogue (TSD 2016).
- S Aswin Shanmugam and Hema A Murthy, "A Hybrid Approach to Segmentation of Speech Using Group Delay Processing and HMM Based Embedded Reestimation", INTERSPEECH 2014.
- Arun Baby, Jeena J Prakash, Rupak Vignesh and Hema A Murthy, "Deep Learning Techniques in Tandem with Signal Processing Cues for Phonetic Segmentation for Text to Speech Synthesis in Indian Languages". pp. 3817 - 3821. 10.21437/Interspeech.2017-666.

Acknowledgement

The authors would like to thank MeitY, Govt. of India for funding the project - "Development of Text-to-Speech synthesis for Indian languages".