MUCS 2021: MUltilingual and Code-Switching ASR challenges for low resource Indian languages

Anuj Diwan¹, Rakesh Vaideeswaran², Sanket Shah³, Ankita Singh¹, Srinivasa Raghavan⁴, Shreya Khare⁵, Vinit Unni¹, Saurabh Vyas⁴, Akash Rajpuria⁴, Chiranjeevi Yarra⁶, Ashish Mittal⁵, Prasanta Kumar Ghosh², Preethi Jyothi¹, Kalika Bali³, Vivek Seshadri³, Sunayana Sitaram³, Samarth Bharadwaj⁵, Jai Nanavati⁴, Raoul Nanavati⁴, Karthik Sankaranarayanan⁵, Tejaswi Seeram⁷, Basil Abraham⁷

¹Computer Science & Engineering, Indian Institute of Technology (IIT), Bombay, India
²Electrical Engineering, Indian Institute of Science (IISc), Bangalore, India
³Microsoft Research India, Hyderabad, India
⁴Navana Tech India Private Limited, Bangalore, India
⁵IBM Research India, Bangalore, India
⁶Language Technologies Research Center (LTRC), IIIT Hyderabad, India
⁷Microsoft Corporation, Bangalore, India
Most of the languages are low-resource.

Code-switching naturally happens between Indian languages and English.
Introduction

Language diversity

- Most of the languages are low-resource.
- Code-switching naturally happens between Indian languages and English.

Language characteristics

- Evolved from four language families.
- Similar based on demographic regions.
Introduction

Language diversity
- Most of the languages are low-resource.
- Code-switching naturally happens between Indian languages and English.

Language characteristics
- Evolved from four language families.
- Similar based on demographic regions.

Objectives
1. **Multilingual ASR**: Exploring the common acoustic properties for better ASR model.
2. **Code-switching ASR**: Exploit code-switching patterns to help in modelling ASR.
Introduction

Language diversity

- Most of the languages are low-resource.
- Code-switching naturally happens between Indian languages and English.

Language characteristics

- Evolved from four language families.
- Similar based on demographic regions.

Objectives

1. **Multilingual ASR**: Exploring the common acoustic properties for better ASR model.
2. **Code-switching ASR**: Exploit code-switching patterns to help in modelling ASR.
Data details: Multilingual ASR

1

“Indian Language Speech sound Label set (ILSL12)”, Indian Language TTS Consortium and ASR Consortium
Data details: Multilingual ASR

A total of ∼450 hrs.

1, “Indian Language Speech sound Label set (ILSL12)”, Indian Language TTS Consortium and ASR Consortium
Data details: Multilingual ASR

A total of ~450 hrs.

Grapheme set:
- Indian language speech sound label set (ILSL12) standard.
- Size: 69, 61, 68, 64, 50 and 65.

1, “Indian Language Speech sound Label set (ILSL12)”, Indian Language TTS Consortium and ASR Consortium
Data details: Code-switching ASR
Data details: Code-switching ASR

- A total of ~150 hrs.
Data details: Code-switching ASR

- A total of $\sim 150$ hrs.
- Technical content of lectures
  - Sentence time-stamps available
  - Alignment is used to obtain audio files for each sentence segment.