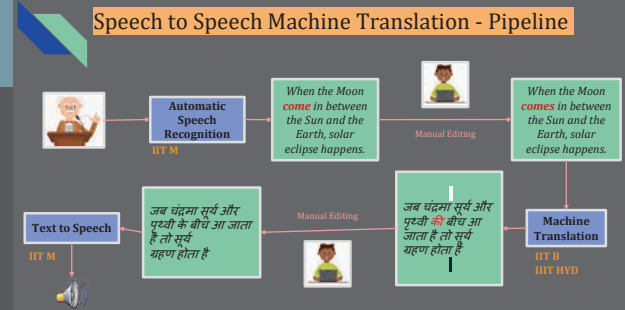


## Bahubhashak: Machine Translation involving English and Indian Languages



**Problem Statement:** India is a land of diversity. The vast number of languages spoken in the country is a testimony to this fact. There are four language families, with twenty-two scheduled languages, with more than thirty languages spoken by over 1 million people. This diversity of languages does bring with it a set of challenging tasks. One of them is education, the primary concern being enabling learning in Indian languages. Teaching-learning in one's mother tongue is known to be highly effective. Responding to this challenge, a consortium of institutes consisting of IIT Bombay, IIT Madras and IIIT Hyderabad has created the speech-to-speech machine translation (SSMT) system from English and Hindi to many Indian languages.

**Uniqueness of the Solution:** The researchers aim to build a speech-to-speech machine translation (SSMT) system that will translate speech in English and Hindi into Indian languages,

focusing on educational content. SSMT consists of a pipeline of stages: (i) first the spoken utterance is converted to text (ASR), (ii) then the produced text is translated to the target language text (MT), and (iii) finally, the translated text is rendered into speech (TTS) in the target language. The uniqueness of the system is it can (a) cater to a large number of Indian languages (b) provide high-quality output that overcomes many unique challenges of Indian languages like phonetic, morphological, syntactic complexity, code-mixing, data sparsity and others.

**Current Status of Technology:** A prototype is developed, which is ready for production/deployment.

**Societal Impact:** A system called "AskAgastya", which is a solution to enable children in rural areas to clarify doubts over a phone line has been developed as a prototype. Ask in the local language (ASR), translates and queries

English textbooks, Respond to students in vernacular voice over the phone line. The system can have a tremendous impact on enhanced access to learning in rural areas across the country.

**Patent(s):** Nil

**Relevant Industries:** Science and Technology, Education, IT, Governance & Policy, Judicial System, Tourism.

**Faculty:** Prof. Pushpak Bhattacharyya, Computer Science & Engineering.