Concept Paper for Voice /Speech based Information Access in Indian Languages on Mobile and Wireless Devices

Technology Development for Indian Languages Programme(TDIL)
Dept of Electronics & Information Technology
Ministry of Communications & Information Technology
Govt. of India
Concept Paper for Voice based Information Access in Indian Languages on Mobile and Wireless Devices

1. Background

Voice based technologies such as Speech Recognition, Text to Speech and Speaker identification are gaining prominence with the advent of next generation Mobile Communication Network namely 3GPP2 and 4G LTE. The industry leaders in the area of Mobile and Voice technologies are building innovative applications and services both at device level and in distributed environment like Cloud based architecture. The accelerated usage of Mobile Device for internet for Browsing and search enhances the potential of speech technologies for Mobile ecosystem.

The mobile and wireless user-base in India is growing in India at a very rapid pace, far exceeding the PC penetration in India, and it is fast becoming a ubiquitous platform for Information access in India. However, the challenge in India is manifold comparison to the European and US scenario. India having 22 Constitutionally recognized languages, 12 Scripts and numerous dialectical variation with complex mapping between scripts and languages, the problem of enabling Mobile Ecosystem in India and their seamless interoperability across devices and platforms makes the task gigantic. Additionally, to proliferate the benefits of ICT technologies the huge non-English Speaking population in India and access of Citizen Centric Services, it is essential that Information Access should be provided in users own language. Voice interface and speech based interface on Mobile platforms would be the key enable technology to achieve the above goal. The recent notification of M-Governance Framework for implementation of e-Gov services through mobile devices conforms to the above objective.

In India, the present research and development efforts to evolve innovative solutions in the speech and Mobile technology is still in a very nascent stage and it is thus necessary that concerted effort needs to be initiated involving all stake holders from academia, industry, standardization and regulatory bodies.
2. Present State of the Art for Speech and Mobile Technology in India:

2.1 Speech:

Speech Technology Research is still in a very nascent stage in India compared to the EU and US Counterparts.

The following is the status of speech research in India:

(1) Speech Resources:

- 50 Hours of Annotated Speech Corpora for 8 Indian Languages, Hindi, Bengali, Assamese, Marathi, Punjabi, Tamil, Malayalam, and Bodo has been developed.
- Preliminary work on speech technology standards such as International Phonetic Alphabet representation for few Indian Languages have been initiated.
- Development for Pronunciation Lexicon for 7 Indian Languages as per W3C Standard has been initiated in consortium mode.

(2) Speech Technology / Applications

- Text to Speech (TTS) in Indian Languages
  Text to Speech (TTS) integrated with screen reader for six Indian Languages have been developed and released in public domain at TDIL Data Centre (http://www.tdil-dc.in) for user feedback and improvement. Approximately 500 visually challenged persons have been trained at various regional centres of National Association of Blind (NAB).

- Automatic Speech recognition (ASR) in Indian Languages
  Alpha version of Automatic Speech Recognitions (ASR) in Indian Languages developed for Agricultural commodity prices information. Independent testing and evaluation is underway for deployment as Voice Interface on NIC Agmarknet Server (http://www.agmarknet.nic.in).

- Very Few Indian industries actually develop commercial applications of TTS and ASR. In Indian Languages
(3) Speech Technology Standards

The Standardization Work for Speech Technology in India is almost Non-existent. Preliminary work has been initiated for standardization of Phonetic Representation of Indian languages. Work is also being initiated for standardization and incorporation of Indian Languages requirements in major Speech standards like W3C –PLS, SSML, SGRS, VXML and HTML Speech and IETF MRCP 2.0.

2.2 Mobile

Though the Mobile user base in India has reached over 900 Million, thereby improving its teledensity, the access to Mobile in Indian Languages still limited and fragmented. Though few of the Mobile device manufactures and Network providers supports Indian languages, the seamless user experience is still missing and lot of research and development activities need to be carried out in this direction.

Some of the seeding activities that are being implemented are:

1. Mobile Font and shaping engine development under National Roll Out Plan project of TDIL by C-DAC Pune

2. Development of Text to Speech for Indian Languages for Android Platform in six Indian Languages by IIT Chennai.
3. **Statement of Objective:**

   *Development of Voice-based information access for Mobile and Wireless Hand-held devices in Indian Languages.*

4. **Project Goal:**

   (1) To explore adaptation of the existing technologies developed under the TDIL Programme and possible augmentation for Mobile and Speech Technologies and development of new technologies.

   (2) To develop and deploy the developed technologies for voice based interface for citizen centric e-applications especially under the National E-Governance Framework using national cloud infrastructure.

   (3) To carry out long term research to incubate innovation and new breakthroughs in the speech and Mobile technology domain and their implementation in Indian Languages.

   (4) To address the standardization needs of fast changing mobile ecosystem and web based mobile access requirements.
5. Implementation Details

Technology Architecture

Fig1. Mobile and Speech Interface and its implementation

All the Technologies and Resources that have been developed so far under the TDIL Programme would be integrated in the Present Framework.
5.1 Technology, Resources and Standard for Speech Technology Domain:

1) **Speech Resources:**

- Development of Annotated Speech Corpora for Min. 300 hours each for all 22 Indian Languages.
- Development of Annotated Speech Corpus for Desktop, Mobile, Wireless and In-vehicle Environments.
- Development of Reference Phoneme set for Indian Languages.
- Development of Speech Recognition Grammar for all Indian Languages.
- Development of Pronunciation Lexicon (PLS) for both Speech recognition and speech synthesis applications in Indian Languages as per W3C Standards.

2) **Applications Development**

- Text-to-Speech in Indian Languages (TTS) integrated with screen reader, MSOffice/Open Office, Web and e-mail clients having average Mean Opinion Score (MOS) 3.9 and above.
- Development of Domain specific Low and Medium Vocabulary Automatic Speech recognition (LV ASR and MV ASR) in Indic Languages.
- The domains are innovative applications in Mobile, E-Governance, agriculture, Healthcare, Railways and Geo-spatial information system.
- Development of Continuous Speech Recognition (CSR) in Indic Languages to develop dictation systems.
- Phonetic search engine in Indian Languages having applications of Word-Spotting from real time continuous speech and topic search for Voice based search engine etc.

3) **Long–Term Research:**

- Distributed Speech Recognition for Next generation Mobile and Wireless Systems [3GPP2 and 4G LTE] for Indic languages
- Voice Browser in Indian languages
- New Models for improvement of performance of TTS in Indian Languages and Implementation of various Signal Processing algorithms for improvement of performance specific to Indian Languages requirements.
- New Models for Speech Recognition in Indian Languages
- Experimental Study of Emotion Recognition.
- Prosodic Modeling and features study.
Speech Technology Standards:

- International Phonetic Alphabet (IPA) in Indic languages
- Speech Corpora for Indic languages
- Internationalization of W3C PLS 2.0 standard for Indian Languages
- Internationalization with respect to Indian Languages requirements in various other W3C Speech Technology standards e.g. SSML 1.1, SGRS, VXML 3.0, EMMA 1.1 and HTML–speech.
5.2 Technology, Resources and Standard for Mobile Technology Domain:

(1) Mobile and Wireless Devices
- Development and Standardization both 9 key Keypad and Qwerty Keypad in Indian Languages.
- Development of Mobile Font in Indian Languages
- Development Mobile Rendering and Shaping Engine in Indian Languages and their support across devices and platforms
- Text Editing Tool, Handwriting and Recognition Tools in Indian Languages
- English to Indian Languages Dictionaries in Indian Languages, complete Localization of User Interface for Mobile Devices.

(2) Transmission and Networks
- Development of Transmission and Reception Protocol and Their Standardization for both Messaging and Internet Applications and their backward compatibility
- Compression techniques for seamless transmission of Indic Language Characters across different networks e.g. 2G, 3G and 4G LTE.
- Integration and Deployment of Mobile Cloud based services in Indian Languages

(3) Mobile Web and Content
- Localization of Mobile Web Browser in Indian Languages
- Development of Mobile Web Content following W3C Mobile Web Best Practices in Indian Languages
In the present Mobile Ecosystem, Multiple Standards exists and Indian Language requirements in these standards are not properly represented. It is thus necessary that, Mobile Messaging and Web standardization w.r.t. Indian Languages necessary for seamless user experience.

Some of the Important Standards that need to be developed quickly are

1. SMS & MMS Standard in Indian Language

2. Implementation / Modification of W3C Mobile Web Standards in Indian Languages.
5.3 Testing and Evaluation for Speech Technologies & Resources

Testing and Evaluation Framework as per international Standard such as NIST and EU need to be developed.

(1) Creation of Testing and Evaluation Framework based on international standard incorporating Indian Languages requirements.
(2) National Level testing and evaluation by independent agencies like STQC and BIS.
(3) Creation of Test-Data and organization of Testing and Evaluation Workshops in the line of Blizzard Challenges.

6. International Collaboration:
   (1) Initiation International Collaboration including collaborative research with USA, EU and ASEAN Countries in the area of Speech and Mobile Technology.
   (2) Signing of Country Level MOUs with above countries to incubate the research and next generation speech technology in Indian Languages.

7. Duration of the Programme
   5 years

8. Proposed Budget Outlay: Total Rs 100 Crore
   National Level Mission Mode Project Proposals in the areas of (i) Speech Technology and (ii) Mobile Technology Development in Indian Languages each having budget Outlay of Rs. 50 Crore each.

9. Expected Impact:

   The Development of Mobile and Speech Technology in Indian Languages is expected to create the next big wave of ICT empowerment to our citizens, through which user can access information and share knowledge and thoughts in their own languages to build India as a True Knowledge Society.