

# Instructions to Participants

\*\*\*\*\*

Since we are having hands-on session for building Isolated and Continuous speech recognition systems using Sphinx and HTK, all the participants are requested to collect & bring the recorded speech data. You can record data in your native language which may be helpful in your further work.

**Software for recording:** Audacity is free software available on net <http://audacity.sourceforge.net/>  
If you have any other software, you can use it.

**Recording specifications:** Sampling rate: 16KHz, bits: 16

**File extension:** should be in wav (PCM)

**Isolated word speech recognition data:**

\*\*\*\*\*

- 1) Minimum of 200 words spoken by four speakers
- 2) Fifty words for each file that means you have 4 files for each speaker
- 3) Each word should have 3 to 4 seconds gap

**Continuous speech data:**

\*\*\*\*\*

- 1) 100 sentences where each sentence is more than 5 words
- 2) sentences should be spoken fluently
- 3) 25 sentences for each file that means you have 4 files for each speaker
- 4) Each sentence should have 6 to 8 seconds gap between them
- 5) Minimum of four speakers

Please also go through speech recognition course slides and other topics on following links

\*\*\*\*\*

<http://course.sls-atl.com/?q=asr/2010/2>

[http://www.abelard.org/briefings/phonetic\\_chart\\_british\\_english.php](http://www.abelard.org/briefings/phonetic_chart_british_english.php)

[http://www.acoustics.hut.fi/publications/files/theses/lemmetty\\_mst/chap3.html](http://www.acoustics.hut.fi/publications/files/theses/lemmetty_mst/chap3.html)

If you have the following book, please go through second chapter for speech production and phonetics basics

“Fundamentals Of Speech Recognition”, by Lawrence Rabiner, B. H. Juang , Biing-hwang Juang

**Note: Please try to learn basic linux commands and one programming language (best is perl) before coming here.**