

PAN localization Project Phase II

Mongolian Speech Recognition

Prof. Damdinsuren Bayanduuren President of MUST, Mongolia
22 – 27 January, 2007 Thimphu, Buthan

Some brief information about MUST

MUST - Mongolian University of Science and Technology.
MUST is the largest state university of Mongolia which offers training in of engineering and technology fields. MUST has 17 schools and 40 research institutes and centers. More than 20,000 students. Up to 1,000 faculty members. 13 schools are situated in capital city Ulaanbaatar, other 4 schools are located in provincial centers as Darkhan, Erdenet, Uvurkhangai, Baruun-Urt. Among 17 schools are two schools related to IT as Atal Bihari Vajpyee School of Telecommunications and Information Technology(STIT, School of Computer Engineering and Management. In these two schools work 80 ICT professionals. At the STIT established ICT Training Center of Excellence including 3 main labs: Microsoft Lab, Cisco engineering Lab, Java Lab, 10 CICs in provincial centers, Videoconferencing system between STIT&NIC(India). Also at STIT established Distance Learning Network over the Mongolia, to which connected all cities and provincial centers. The system provides real time interactive audio-video connection and At cities and provincial centers created 25 centers with 25-30 seats each. Therefore at once can be involved more than 700 students. LMS as Moodle is in use at STIT. Translated ICT terminologies and adopted as standard. E-school program is being implemented at MUST(STIT). But we need some financial support for these purposes.

Mongolian language

Mongolian is a Mongolic language. The Altaic theory proposes that the Mongolic family is a member of the larger Altaic family. Khalkha Mongolian is the national language of Mongolia. Over 2.7 million people speak Mongolian throughout Mongolia.

Mongolian language: Lexicon

In Mongolian, there are loan words especially from Old Turkic, Sanskrit (often through Uigur), Tibetan, Tungusic and Chinese, and, more recently, from Russian, Chinese and English. Commissions in Mongolia have been busy by translation of new terminologies into Mongolian.

Mongolian language: Phonology

Genuine Mongolian words use the following syllable structures:

V, CV, CVC, VVC, CVCC, VCCC, CVCCC, Syllablification is $L \rightarrow R$.

The long vowel also happens to be stressed, since in standard Mongolian stress falls on the first long-vowel led syllable of a word. If there is no long vowel, the word-initial syllable is stressed.

Mongolian language: Morphology

Modern Mongolian is an agglutinative, exclusively suffixing language the suffixes of which are most often composed of a single morpheme. It has a rich number of morphemes to build up more complex words from simple roots. Mongolian language syntax has SOV structure.

Mongolian Language has about 1000 diphones. From these a diphone lexicon of about 600 diphones has been produced. Research on NLP (Natural Language Processing) is also being done.

There are approximately 40000 – 50000 words in Mongolian language. There are about 10000 root words, and about 300 suffixes.

Research background

On speech signal processing research works are conducted in two directions as speech synthesis and speech recognition. Doctorial students supervised by me defended their dissertations on speech synthesis: 1. By Dr. B. Otgonbayar done synthesis of vowels. Caried out mathematical models of vowels; 2. By Dr. O. Bat-Enkh done synthesis of words and sentences. Used approach based on modulation theory; 3. By Dr. Ts. Baatarkhuu done investigation on information and statistical characteristics of Mongolian language. Investigated probabilities of occurance of 1 chatacter, 2, 3, 4, 5, 6, 7, 8 characters in Mongolian speech and text materials; Now 2 doctarial students are doing their research works on diphone synthesis and word root based synhtesis.

Domain-specific synthesis research project has been done in 2005 for visually challenged people, with over one thousand words in the domain. Speech synthesis - the artificial production of human speech - has inspired an enormous body of research.

Concatenative synthesis approach has been chosen as the main speech synthesis method.

Diphone synthesis research project is currently in progress. Word root based TTS also in process.

Done research work funded by Foundation of Science and Technology of Mongolia. In the frame of this work conducted some spectrum analysis of Mongolian sounds for male and female and different ages. Actively participated Dr. Z. Buyankhishig, G.Demberel and determined formants.

In the summer of 2006, two researchers have attended the Summer School on Asian Language Processing in Pakistan.

The project which will be done near 2.5 year funded by IDRC.

Mongolian Speech Recognition

- We have developed a project in accordance with the Pan localization program.
- Project team will consists of a project leader, faculty advisor, three research officers, senior linguist advisor, and two research assistants.
- Project will be started on May 16, 2007 and continue for 30 months.
- Training of team will be conducted before the project starts

Project has following phases

- **Evaluation of Toolkits**
 - Selection of Toolkits
 - Selection of Test Vocabulary 100 words
 - Develop speech corpus for selected vocabulary
 - Evaluate Toolkit1
 - Set up Toolkit 1
 - Add vocabulary for training
 - Test Performance
 - Evaluate Toolkit2
 - Set up Toolkit 2
 - Add vocabulary for training
 - Test Performance
 - Research report on toolkit evaluation and selection
- **Training of selected toolkit**
- **Creation of vocabulary**
- **Integration and testing**

After completion of this project we plan to start working on other related topics.

Thank you for attention.