The Third International Conference "Problems of Cybernetics and Informatics"
September 6-8, 2010, Baku, Azerbaijan. Section #1 "Information and Communication Technologies"
www.pci2010.science.az/1/04.pdf

HMM-BASED LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION SYSTEM FOR AZERBAIJANI

Ali Abbasov[1], Rauf Fatullayev[2], Abulfat Fatullayev[3]

[1,2]Ministry of Communications and Information Technologies of AR, Baku, Azerbaijan
[3]Cybernetics Institute of ANAS, Baku, Azerbaijan
[1]*ali@ab.az*, [2]*fatullayev@gmail.com*, [3]*fabo@box.az*

Abstract. This article describes the main stages of the development and characteristics of a speech recognition system for the Azerbaijani language developed within the project Dilmanc. Some numerical data on the amount of information to create the database for normal functioning of the developed software modules and some characteristics of a speech recognition system for the Azerbaijani language are given.

Key words. ASR systems, HMM-based speech recognition, the Azerbaijani language, Turkic languages

Introduction

Turkic group of languages is one of the largest groups comprising more than 50 languages (Azerbaijani, Kazakh, Kyrgyz, Tatar, modern Turkish, Turkmen, Uzbek, etc.). Nearly 40 of these languages at the moment has a certain range of existence, 15 of these languages are dead languages [1]. Despite the impressive number, most of these languages (except modern Turkish language) are little investigated languages. Language processing technologies (automatic spell and grammar checking, speech understanding, recognition and synthesis, machine translation and etc.) and researches on the development of such systems are very rare for these languages.

Azerbaijani language as one of the Turkic languages is also a little studied language in terms of creation of modern language technology applications, despite a certain amount of research conducted in the 80th years of last century [2-3].

To fill this gap, the project *Dilmanc* was initiated within the national e-governance project. Concrete goals for the development and implementation of formal linguistic technologies for the Azerbaijani language were set for this project. The project was initiated at the beginning of 2005 and continues till now. The project is supported by the Ministry of Communications and Information Technologies of the Republic of Azerbaijan and the UN Development Programme (UNDP-Azerbaijan).

Research works within project are being conducted in different directions, some of which are presented below:

- Machine translation from European languages into the Azerbaijani language and vice versa;
- Machine translation between Turkic languages;
- Speech Recognition for the Azerbaijani language;
- Speech synthesis for the Azerbaijani language;
- Speech to speech translation and etc.

It should be noted that this article describes finished modules of software products, information about what and how has been done can be found in the publications of the developers group's members [4-8].

The Third International Conference "Problems of Cybernetics and Informatics"
September 6-8, 2010, Baku, Azerbaijan. Section #1 "Information and Communication Technologies"
www.pci2010.science.az/1/04.pdf

Azerbaijani speech recognition system DilmancASR

Speech recognition system for the Azerbaijani language is based on Hidden Markov Models (HMM) and is designed for continuous speech recognition for the Azerbaijani language (Fig. 2).

Fig. 1. Azerbaijani speech recognition system

The idea of using HMM for recognition of human speech is not new and was proposed long ago. The first such computer system (based on HMM) was developed in IBM Corporation [9]. It should be noted that apart from HMM, there are other approaches to the development of human speech recognition systems [10]. But the most successful speech recognition systems are developed on the basis of HMM. Examples include speech recognition module of the Windows operating system and a very popular speech recognition system Dragon NaturallySpeaking (http://www.nuance.com/naturallyspeaking).

Azerbaijani researchers conducted research on the development of speech recognition system for the Azerbaijani language based on artificial neural networks [11], but these attempts didn't yield tangible results.

DilmancASR system is yet the first and the only system designed for speech recognition for the Azerbaijani language. The system can be used for dictation of text in any text editor (MS Word, WordPad, etc.). The system allows you to configure the microphone based on the speaker's speech amplitude and noise environment (Figure 2), as well as to determine the quality of recording of microphones, because these parameters are critical for recognition accuracy.

DilmancASR (as well as other speech recognition systems based on HMM) consists of two blocks:

1. acoustic recognition module;
2. language module for correction.

For the acoustic recognition module, following information bases have been developed and are being used:

- Speech corpus of about 60 hours' duration;
- The speech corpus contains the voice of over 1500 speakers;
- Over 40,000 prompts containing the most frequently used diphones and triphones of Azerbaijani language have been dictated and recorded;
- Nearly 15,000 recorded prompts have been annotated at the phoneme level;
- Allophones database containing over 500,000 items has been developed;
- A database of about 150 thousand most frequently used word forms of the Azerbaijani language;

Linguistic correction module consists of 1-4-grams and their number is over 26 million. These 4-grams are based on the text corpus, which consists of more than 300 million word forms. It should be noted that for the low density languages (a group of languages with limited electronic resources, to which so far Azerbaijani language belongs), the development of large corpora is also quite difficult.

24

The Third International Conference "Problems of Cybernetics and Informatics"
September 6-8, 2010, Baku, Azerbaijan. Section #1 "Information and Communication Technologies"
www.pci2010.science.az/1/04.pdf

Information databases designed for acoustic and linguistic blocks allow recognizing continuous speech in the Azerbaijani language accurately enough. The average accuracy of recognition is about 80%, but on individual subjects, recognition accuracy reaches 90%. Works for the improvement of the system are being continued. Since the recognition accuracy depends on the amount of speech corpus and allophone database, works to increase these tools (Fig. 2 and 3) and other information databases are going on.
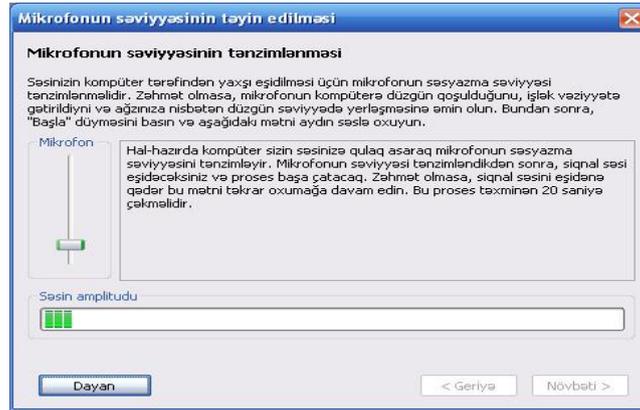

Fig. 2. Microphone setup
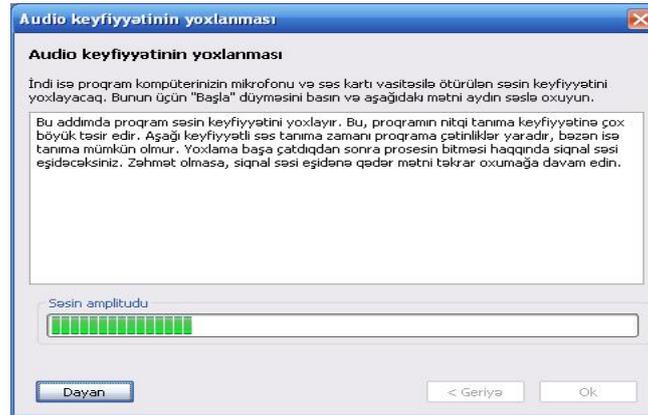
Fig. 3. Testing the quality of the microphone

Conclusion

DilmancASR is the first and the only speech recognition system in the Azerbaijani language. Currently, the program module for adapting the system to a specific user is being developed. After connecting the module to the system recognition accuracy will certainly increase.

Researches within Dilmanc project on other areas (improvement of machine translation systems, telephone translator, speech synthesis and etc.) are currently going on. For all the above areas, project team has got practically applicable results and works on the development and improvement of software products on the basis of these results are being continued.

The Third International Conference "Problems of Cybernetics and Informatics"
September 6-8, 2010, Baku, Azerbaijan. Section #1 "Information and Communication Technologies"
www.pci2010.science.az/1/04.pdf

References

1. Kibrik A.A., Tenishev E.R., Poceluevskij E.A. and Kormushin I.V. 1997. Jazyki mira: Tjurkskie jazyki [Languages of the world: Turkic languages]. Moscow: Indrik. 542 p

2. Махмудов М.А. Инженерно-лингвистические исследования в тюркском языкознании / Всесоюзная конференция «Переработка текста методами инженерной лингвистики», Минск, 1982, с. 17-19

3. Махмудов М.А. Лексико-морфологический МП азербайджанского текста на русский язык / Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Чимкент, 1980, с. 257-260

4. Fatullayev R, Abbasov A, Fatullayev A. 2008. Dilmanc is the 1st MT system for Azerbaijani. In: Proc. of SLTC-08, Stockholm, Sweden, pp.63-64

5. Fatullayev R, Abbasov A, Fatullayev A. 2008. Peculiarities of the development of the dictionary for the MT System from Azerbaijani. In: Proc. of EAMT-08, Hamburg, Germany, pp.35-40

6. Fatullayev R, Mammadova S, Fatullayev A. Translating composite sentences in Azerbaijani-English MT system. ISMTCL: International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages, and their application to emergencies and safety critical domains, July 1-3, 2009, Besançon, France, pp.104-113

7. Fatullayev R, Shagavatov S, Fatullayev A. Predicate-Subject agreement in Azerbaijani and its presentation in an MT system. 15th International Conference on Soft Computing (Mendel 2009), June 24-26, Brno, Czech Republic, pp. 309-314

8. Fatullayev R, Abbasov A, Fatullayev A. Overcoming Agglutination Difficulties in the development of an MT system from the Azerbaijani language. Speech and Language Technology. International journal of Polish phonetics association, Poznan, vol. 11 (2008), pp. 215-228

9. Rabiner, L.R., "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. of IEEE, 1989, 77(2), pp. 257-286.

10. A.Weibel, T.Hanazawa, G.Hinton, K.Shikano and K.Lang, "Phoneme Recognition Using Time-Delay Neural Networks," IEEE Trans. Acoustics, Speech, Signal Proc., 37: 393-404, 1989

11. Aida-zade K.R., Ardil C., Rustamov S.S. Investigation of Combined use of MFCC and LPC Features in Speech Recognition Systems. IJSP "International Journal of Signal Processing", Volume 3, 2006, pp. 105-111.