Russian Speech Recognition for Telecommunications

Andrey Ronzhin, Alexey Karpov, Izolda Li
St.Petersburg institute for informatics and automation of RAS
39, 14 Line, 199178, St. Petersburg, Russia
E-mail: {ronzhin,karpov,lee}@iias.spb.su Web: www.spiiras.nw.ru/speech

Introduction

The development of new services, which could maximally use natural modalities of a human and, first of all, the natural speech is a very important task.

In contrast to English the Russian language has much more variety on word-form level and so the size of recognized vocabulary sharply increases as well as quality and speed of the processing decrease.

The developed speech recognition system SIRIUS (SPIIRAS Interface for Recognition and Integral Understanding of Speech) allows significantly reducing the size of vocabulary and time for speech processing due to additional (morphemic) level of speech representation.

Automation of operator's functions

Control systems for information streams on the basis of the call-centers of a various origin (automation of operator's functions);

Telephone services, including a mobile content (automation of directory systems);

Automated inquiry machine located in hotels, airports, train stations, subway, supermarket, tourist places.

Call centres market in Russia

The main cost of call centers is a salary of operators. It is over 70% of whole costs.

The processing of 1 minute of incoming call by alive operator costs 0,07$-0,12$.

The cost of automatic call processing does not exceed 0,02$ per 1 minute.

By 2007 about 15000 telephone lines will be automatically served by speech recognition systems.

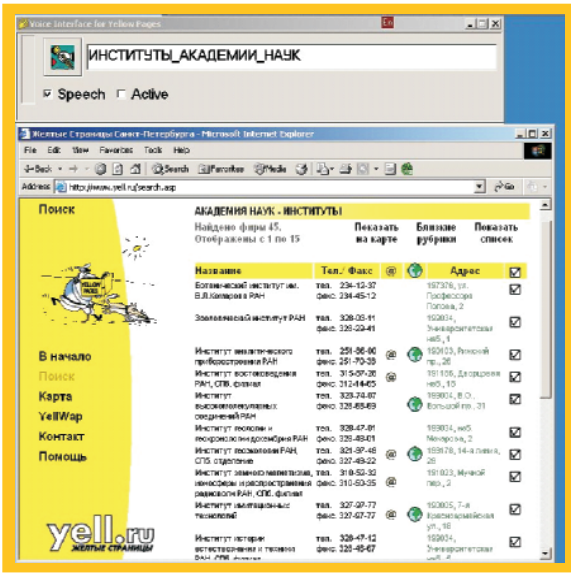
Volume of a segment of the call-centers will make about 7,5 M Euro.

Advantages of our approach

Quality and robustness of Russian speech recognition. The unique algorithm of speech recognition, specially constructed in view of morphological features of Russian is developed.

Geographical closeness to consumers. Involving the developers in interaction with consumers allows to construct a qualitative feedback with the market, trace reaction of consumers, change of language and improve quality of the product.

Presence of a technological platform for mass testing and training the system directly in the product consumption. New Voice Telecom Russia

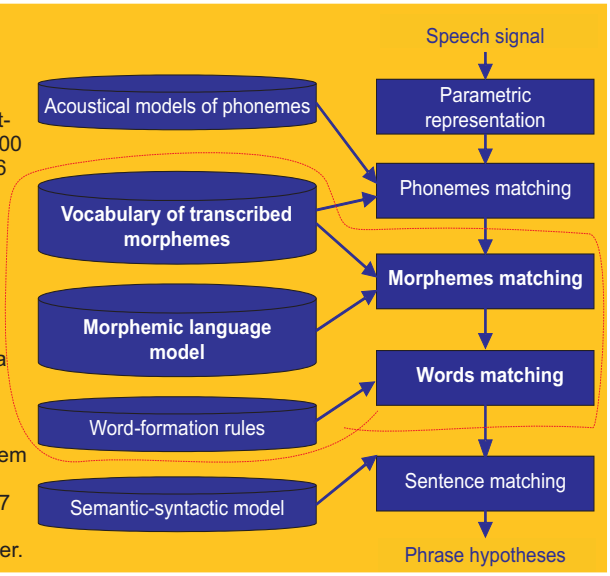


Experiments

Voice access to the electronic catalogue “Yellow Pages of Saint-Petersburg” (about 2000 rubrics, 5 speakers, 16 kHz, accuracy over 95%).

The automatic telecommunication service “Easy Conferencing” (recognition of the telephone numbers via telephone line, 50 speakers, 8kHz, accuracy over 92%).

The developed system based on morpheme recognizer works in 1,7 times faster then the words-based recognizer.



Conclusion

The conducted research is directed to the investigation of peculiarities of Russian speech and the creation of large vocabulary speech recognition.

The voice access to Yellow pages is being developed and tested now. Besides, the system was tested by telephone line via usual modems and specialized telecommunication plats Intel Dialogic.

The future work in this research will consist in the gathering the large databases with Russian telephone speech. Experimental exploitation of telephone service including voice access to "Yellow Pages of Saint-Petersburg" will allow accumulating large speech database and realize optimal dialogue strategy.

Acknowledgement

This work is supported by the Saint-Petersburg Government, Russian Science Support Foundation, INTAS Project #04-77-7404, SIMILAR NoE FP6:IST-2002-507609 and NewVoice company.

References

A.A. Karpov, A.L. Ronzhin. "Speech Interface for Internet Service Yellow Pages". Intelligent Information Processing and Web Mining: Advances in Soft Computing, Proc. of the International IIS: IIPWM’05 Conference, Gdansk, Poland, Springer-Verlag, 2005, pp. 219-228.

A.L. Ronzhin, A.A. Karpov. Large Vocabulary Automatic Speech Recognition for Russian Language. In Proc. of Second Baltic Conference on Human Language Technologies, Tallinn, Estonia, 2005, pp. 329-334.

A.L. Ronzhin, A.A. Karpov. Implementation of morphemic analysis for Russian speech recognition, 9-th International Conference SPECOM’2004, St. Petersburg: “Anatoliya”, 2004, pp. 291-296.