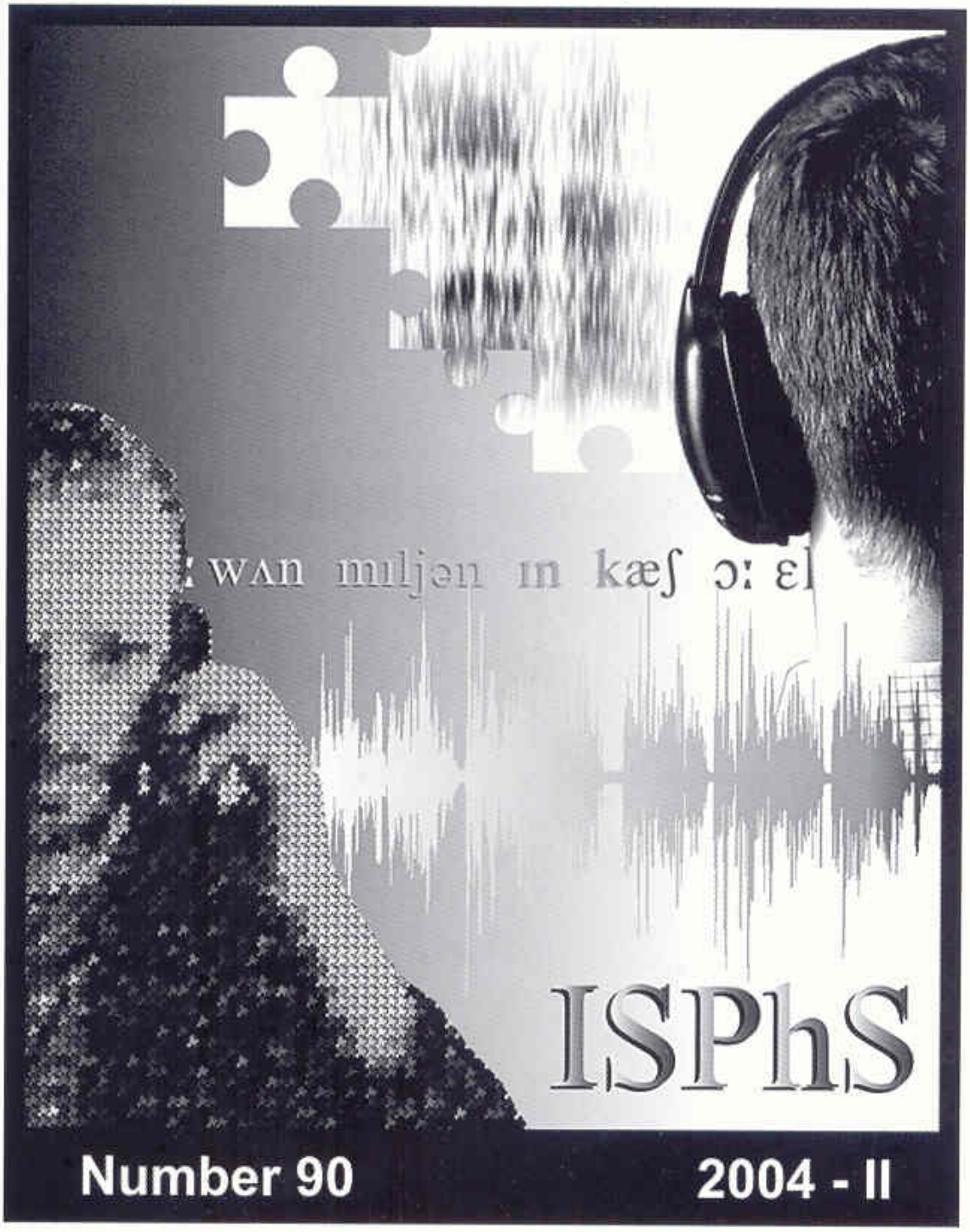
the Phonetician

A Publication of ISPhS/**International Society of Phonetic Sciences**



Number 90 **2004 - II**

The original version of the paper is published: The Phonetician. A publication of the International Society of Phonetic Sciences, Vol. 90, 2004, pp. 34-38.

Conference Reports

9th International Conference SPECOM September 20-22, 2004, St. Petersburg, Russia

During September 20 through September 22 2004 the 9th International Conference "Speech and Computer" was held in Saint-Petersburg. The series of Conferences SPECOM was organized in 1996 by Saint-Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS) in collaboration with Hertsen Pedagogical University. The first Conference took place in Saint-Petersburg, then in 1997 in Romania, 1998, 2000, 2003 – in Saint-Petersburg, 1999, 2001, 2003 – in Moscow. This year the Conference was especially successful and over 170 persons from 36 countries took part at the Conference during 3 days.

The interaction between a human and a computer in natural form, like conversation between humans, is one of the most important and complex tasks of Artificial Intelligence. The existent models of speech recognition and understanding are essentially worse in comparison with human's speech abilities. It is evidence of their insufficient adequacy that restricts the usage of speech technologies both for industry and in private life. At present the active studies are fulfilled in the area of multimodal interfaces, which combine diverse kinds of information input (speech, lips articulation, gestures, gaze direction etc.) in joint system. The multimodal interfaces are absolutely natural for communication between humans. We can select most suitable communicational channels for transmission of any kinds of information in any time of conversation. Such interfaces allow to provide the most effective and ergonomic way of interaction between a human and diverse automatic devices for control and communication.

Conference SPECOM is only one of the few Russian conferences, devoted to speech technologies, which brings together so many foreign participants and has really International character. In many respects it is connected with the scientific themes and questions of the conference (speech, natural language, human-computer interaction). The main objective of these research is creation of effective of natural means for human-human and human-computer interaction in spite of some technical, linguistic and social difficulties. Such conference is extremely important both for scientists and researchers from Russia and New Independent States and for foreign researchers (from Europe, America and Asia) because the conference is a bridge between Russian and International science. The conference is good possibility for Russian scientists to find out about leading achievements in the world science and foreign scientists can know more about studies, which are performed in the New Independent States.

The main aim of all the conferences SPECOM is discussion of the numerous important problems of human-computer interaction and artificial intelligence: automatic speech recognition and understanding, speech synthesis, speaker

verification and identification, etc. Each year the novel developments and results of research in this scientific area are presented at the SPECOM.

The 9th International Conference SPECOM'2004 was organized by Speech Informatics Group of SPIIRAS by financial support of INTAS and SIMILAR European Network of Excellence. The technical program includes large spectrum of novel contributions of 276 authors from 35 countries (NIS, Europe, America, Asia, Australia and Africa.), arranged in 12 oral sessions. 130 papers were published in Proceedings of SPECOM'2004 (which contains over 750 pages) and in electronic version of Proceedings on CD. The conference web site, which includes the information about all previous conferences, is http://www.spiiras.nw.ru/speech/. Some statistics of contributions per country is presented in Table 1.

TABLE 1. NUMBER OF CONTRIBUTIONS PER COUNTRY

53	Russia	2	India
10	Germany	2	Israel
9	Belarus	2	New Zealand
5	Belgium	2	Poland
5	Greece	2	Portugal
5	Japan	2	Romania
5	Spain	1	Algeria
5	USA	1	Argentina
4	France	1	Armenia
4	UK	1	Australia
3	Czech Republic	1	Brazil
3	Finland	1	Canada
3	Korea	1	China
3	Mexico	1	Iran
3	Serbia and Montenegro	1	Lithuania
3	The Netherlands	1	Singapore
3	Turkey	1	Thailand
3	Ukraine		

In the framework of the conference the several special events will be organized: Multimodal Interfaces Day, Special session "Multimodal services and applications for disabled people" and INTAS Strategic Scientific Workshop "Development of perspective applications of Human-Computer Interaction for Information Society".

The INTAS Strategic Workshop was directed to discussion of the achievements and the problems in the area of human-computer interaction and information technologies, as well as elaboration of some required actions for improvement of cooperation between New Independent States of the former Soviet Union and European countries. For this aim the leading world-known scientists, representatives of diverse scientific, industrial and international funding organizations were invited for participation in the Workshop.

Plenary session was started by Director of SPIIRAS Prof. Rafael Yusupov. He emphasized the importance and actuality of research in the domain of automatic

speech recognition and understanding as well as in human-computer interaction in general. Especially Prof. Yusupov marked the significance of new direction in science and technique – multimodal interfaces, which were presented in Russia first. Then the keynote lectures were given by Prof. Jean-Paul Haton (France), President of British Voice Association John Rubin (UK), Prof. Pavel Skrelin (Russia), representative of France Telecom R&D Prof. Katarina Bartkova (France).

Prof. Jean-Paul Haton presented keynote lecture "Automatic speech recognition: past, present and future" with survey of existing advanced methods for speech recognition and about future of automatic speech recognition. The basic principles of statistical approach to ASR based on Hidden Markov Models were considered in this lecture.

Prof. John Rubin presented keynote lecture with survey "Psychogenic voice disorders in performers: a psychodynamic model" about speech production and psychological aspects of speech perception, as well as problems arising during the process of speech production and methods of recognition of "complicated" speech at the persons with specific needs.

Prof. Pavel Skrelin presented keynote speech "Segment features in different speech styles" about differences of reading (i.e. pre-prepared) and spontaneous speech as well as results of recognition of both types of speech. Basic approaches to recognition of isolated and continuous speech as well as created speech databases, which are necessary for training of the systems of automatic speech recognition, were considered in this lecture.

Prof. Katarina Bartkova presented keynote lecture "Foreign accent processing in automatic speech recognition" about differences in organization of recognition systems for different languages (French, English, German, Spanish, etc.) The phonetic structure of these languages and base phonological rules for representation of acoustic models were presented.

Further work of Conference SPECOM'2004 was held within eleven sessions.

Main questions, which were discussed on Section "Multimodal interfaces", were dedicated to the systems in which for man-machine dialog some natural methods of communication such as speech input, gestures, handwriting etc. can be used jointly. Particular interest was caused by the paper of Prof. Ferran Marques (Spain) and Prof. Benoit Macq (Belgium) devoted to the methods for recognition of gesture and position of human as well as methods for combination of these kinds of input with speech input.

Also on this session an interesting paper of Prof. Slobodan Jovicic (Serbia&Montenegro) "Serbian emotional speech database: design, processing and evaluation" about definition of emotional state and usage of emotional colors of speech for human-computer interaction was presented.

On the session "Speech signal processing" papers concerning mathematical modeling of signal, speech filtration, noise filtering were presented. Among presented papers particular interest was caused by the paper of Prof. Hisao Kuwabara (Japan), which described the influence of individual differences of speakers on acoustic parameters of speech and the variety of these parameters. Also the series of papers on new approaches for speech signal processing based on kernel methods was presented by the group of Prof. Andreas Wendemuth (Germany).

Session "Speech recognition" was the main on the SPECOM'2004 and the most of papers were presented at this session. The oral presentations were devoted to the fundamental problems of automatic speech recognition, as well as different aspects of application of speech technologies. Special attention was paid to the technologies of

recognition of speech on national languages (in particular, Russian). The representatives of such companies as IBM (Dimitry Kanenevsky, USA) and Motorola (Chip Wood, USA) presented their advanced research. Particular interest was caused by the paper of Dr. Andrey Ronzhin (Russia), devoted to the absolutely new method for recognition of Russian speech with use of morphemic level of speech representation. In framework of this session several demo-versions of some applied systems of automatic speech recognition were successfully demonstrated (Chip Wood, USA; Dimitri Kanenevsky, USA; Andrey Ronzhin, Russia; Mais Farkhadov, Russia). These demonstrations caused the interest both leading scientists and potential investors and users.

On the sessions “Speech understanding and natural language processing” and “Dialogue, ontologies and Knowledge presentation” particular interest was caused by papers of Prof. Valery Galunov (Russia) and Irina Kobozeva (Russia). In the paper of Prof. Galunov the attempt of construction of general structure of ontology of speech technology is made. Some examples of elements of ontology in the areas of speech recognition and synthesis are considered. In the paper of Irina Kobozeva the usage of semantic and pragmatic information in spoken discourse was considered.

The Session “Speaker recognition” was devoted to the methods for verification and identification of speaker as well as usage of these biometrical methods in safety systems, intended for warning of possible acts of terrorism. This problem is especial actual today in the world. Some interesting scientific results and applied systems were presented by the group of Juhani Saastamoinen (Finland).

On the session “Speech synthesis” methods for synthesis and generation of speech for different languages were presented. Prof. Edward Shpilewski (Poland), Les Doherty (Australia) and the team of Prof. Boris Lobanov (Belarus) presented the papers and demonstrations of synthesized speech.

Within the framework of the conference the Special Session “Multimodal services and application for disabled people” was organized jointly with SIMILAR European Network of Excellence. This scientific direction is aimed for the creation of perspective assistive systems for disabled people. On this session several papers were presented by representatives of SIMILAR NoE describing joint use of gestures and speech input for blind people (Dimitrios Tzovaras, Greece, et al.). One paper about using of cued speech gestures for phone communication between deaf people was presented by Alice Caplier (France). Also the paper, describing the system for humans with disabilities of hands, developed in SPIIRAS (Alexey Karpov, Alexander Nechaev, Russia), was presented during special session.

INTAS Strategic Scientific Workshop ”Development of perspective applications of Human-Computer Interaction for Information Society” was organized 21-22 September. Two-day Workshop contained 3 sessions: “Strategic Scientific Session”, “Presentations of European Research Projects”, “Fundamentals of Human-Computer Interaction”. During the Workshop the several strategic lectures, connected with existent programs and funds for supporting fundamental research, were given by representatives of diverse funding organizations, as well as the lectures according to further development of scientific schools in NIS and Europe were given by leading scientists. Official representative of INTAS Patrizia Asirelli (Belgium) presented the key lecture about priority directions of activities of Association INTAS and some statistics on completed and current research projects in the area of information technologies. FP6 Multiplayer for North-West region of Russia Alexey Ivanov presented the report about experience of participation of Russian organizations in

Sixth Framework Programme (FP6) for 2002-2006. Official representatives of Russian Foundation for Basic Research (RFBR) Nelly Didenko and Andrey Petrovsky presented the reports about cooperation between INATS and scientific society of Saint-Petersburg as well as about grants, which are proposed by RFBR for supporting of fundamental research for Russian scientists.

Also during INTAS Strategic Scientific Workshop several lectures were presented, which described scientific studies corresponded to priority directions of INTAS. In particular, Prof. Taras Vintsiuk (Ukraine) presented the paper about state-of-the-art speech technologies in Ukraine and about future projects and expected results. Prof. Lev Stankevich (Russia), who was the captain of the team which won the football world championship among the robots RoboCup-2004 in Portugal, gave the lecture about perspectives of development of intellectual robots in Russia and successful experience of participation in RoboCup championships. Several reports were made on Presentations of diverse European research Projects. Prof. Benoit Macq (Belgium) told about SIMILAR Network of Excellence integrating a research task force on multimodal interfaces. Prof. Igor Kotenko and Svetlana Stepanova told about their current research projects and some obtained results according to the European projects supported by EC (FP6) and INTAS correspondingly. Prof. Boris Sokolov and Prof. Alexander Timofeev told about ongoing and planned in future studies in the area o human-computer interaction fulfilled in SPIIRAS.

The banquet will take place Monday evening and an attractive social program is organized before, during, and after the Conference for participants and accompanying persons to experience the attractions of the Saint-Petersburg and surrounding areas (Peterhoff, Pushkin, etc.).

On completion of the Conference the Scientific Committee declared that the 9-th International Conference “Speech and Computer” SPECOM’2004 was very successful. The Conference turned out the greatest and the most representative scientific event in Russia, dedicated to speech technologies. As the main results of work of International Conference SPECOM’2004 in Saint-Petersburg can be marked the following:

- evaluation of achievements and finding the significant undecided questions connected with human-computer interaction;
- improvement of level of cooperation between academic and industrial organizations which work in the area of creation of the perspective systems for communication between a human and a computer;
- the acquaintance of Russian researchers and students with novel scientific results of leading international companies and universities;
- promotion of future joint International projects, summers schools and other scientific events devoted to the problems of human-computer interaction in information society.

It is supposed that the next 10-th International Conference “Speech and Computer” will be organized in 2005 in Greece and SPECOM’2006 will be held again in Saint-Petersburg.

Andrey Ronzhin,
Speech Informatics Group of SPIIRAS
e-mail: ronzhin@iias.spb.su