



INTERSPEECH

2011

Keynotes

27-31 August 2011

Firenze Fiera Conference Center
Florence, Italy



INTERSPEECH

2011

Keynote Sessions

Keynote 1

ISCA Scientific Achievement Medallist for 2011

Julia Hirschberg
Columbia University
Department of Computer Science



“Speaking More Like You: Entrainment in Conversational Speech”

Sun-Ses-K1: Sunday, August 28th, 2011 - 11:00, Auditorium

Chair: Isabel Trancoso

Abstract

When people engage in conversation, they adapt their speaking style of their conversational partner in a number of ways. They have been shown to adopt their interlocutor's way of describing objects and to align their accent, their syntax, their pitch range, and their speaking rate to their partner's -- as well as their gestures. They also adapt their turn-taking style and use of cue phrases to match their partner's. These types of entrainment have been shown to correlate with various measures of task success and dialogue naturalness. While there is considerable evidence for lexical entrainment from laboratory experiments, less is known about other types of acoustic-prosodic and discourse-level entrainment and little work has been done to examine entrainments in multiple modalities for the same dialogue. I will discuss research in entrainment in multiple dimensions on the Columbia Games Corpus and the Switchboard Corpus. Our goal is to understand how the different varieties of entrainment correlate with one another and to determine which types of entrainment will be both useful and feasible to model in Spoken Dialogue Systems..

Presenter

Prof. Julia Hirschberg is professor of Computer Science at Columbia University, working on prosody, emotional speech and dialogue systems. She was editor-in-chief of Computational Linguistics from 1993-2003 and co-editor-in-chief of Speech Communication from 2003-2006. She served on the Executive Board of the Association for Computational Linguistics (ACL) from 1993-2003, on the Permanent Council of International Conference on Spoken Language Processing (ICSLP) since 1996, on the board of the International Speech Communication Association (ISCA) from 1999-2007 (as President 2005-2007), and is currently on the board of the CRA-W. She has been a fellow of the American Association for Artificial Intelligence since 1994 and an ISCA Fellow since 2008. She received an honorary doctorate from KTH in 2007 and the James L. Flanagan Speech and Audio Processing Award in 2011.

Keynote 2

Tom M. Mitchell
E. Fredkin University
Machine Learning Department School of Computer Science
Carnegie Mellon University



“Neural Representations of Word Meanings”

Mon-Ses-K2: Monday, August 29th, 2011 - 8:30, Auditorium
Chair: Giuseppe Riccardi

Abstract

How does the human brain represent meanings of words and pictures in terms of neural activity? This talk will present our research addressing this question, by applying machine learning algorithms to fMRI and MEG brain image data. One line of our research involves training classifiers that identify which word a person is thinking about, based on their observed neural activity. A second line involves training computational models that predict the neural activity associated with arbitrary English words, including words for which we do not yet have brain image data. A third line of work involves examining neural activity at millisecond time resolution during the comprehension of words and phrases.

Presenter

Professor Tom M. Mitchell is a professor in the E. Fredkin University and head of the Machine Learning Department at Carnegie Mellon University. His research interests lie in cognitive neuroscience, machine learning, natural language processing, and artificial intelligence. Mitchell is a member of the US National Academy of Engineering, a Fellow of the American Association for the Advancement of Science (AAAS), and Fellow and Past President of the Association for the Advancement of Artificial Intelligence (AAAI). Mitchell believes the field of machine learning will be the fastest growing branch of computer science during the 21st century.

Keynote 3

Alex 'Sandy' Pentland
MIT - Massachusetts Institute of Technology



“Honest Signals”

Tue-Ses-K3: Tuesday, August 30th, 2011 - 8:30, Auditorium
Chair: Giuseppe Riccardi

Abstract

How did humans coordinate before we had sophisticated language capabilities? Pre-linguistic social species coordinate by signaling, and in particular `honest signals' which actually cause changes in the listener. I will present examples of human behaviors that are likely honest signals, and show that they can be used to predict the outcomes of dyadic interactions (dating, negotiation, trust assessment, etc.) with an average accuracy of 80%. Patterns of signaling also allow accurate identification of social and task roles in small groups, predict task performance in small groups, guide team formation, and understand aspects of organizational performance. These experiments suggest that modern language evolved `on top' of ancient signaling mechanisms, and that today linguistic and signaling mechanisms operate in parallel.

Professor Alex 'Sandy' Pentland is a pioneer in computational social science, organizational engineering, and mobile information systems. He directs the MIT Human Dynamics Lab, developing computational social science and using this new science to guide organizational engineering. He also directs the Media Lab Entrepreneurship Program, spinning off companies to bring MIT technologies into the real world. He is among the most-cited computer scientists in the world. His most recent book is `Honest Signals' published by MIT Press.

Keynote 4

Round Table

“Future and Applications of Speech and Language Technologies for the Good Health of Society”

Wed-Ses-K4: Wednesday, August 31st, 2011 - 8:30, Auditorium

Chair: Giuseppe Riccardi

Gabriele Miceli

Università degli Studi di Trento

Dipartimento di Scienze della Cognizione e della Formazione



“Language disorders: viewpoints on a complex object”

Björn Granström

KTH - Royal Institute of Technology

TMH - Department of Speech, Music and Hearing, Stockholm, Sweden



“Speech technology in (re)habilitation of persons with communication disabilities”

Hiroshi Ishiguro

Osaka University - Dept. of Adaptive Machine Systems

Intelligent Robotics Laboratory

Dept. of Systems Innovation

Graduate School of Engineering Science

Visiting Group Leader

Dept. of Communication Robots

ATR Intelligent Robotics and Communication Laboratories



“From teleoperated androids to cellphones as surrogates”

Gabriele Miceli
Università degli Studi di Trento
Dipartimento di Scienze della Cognizione e della Formazione

“Language disorders: viewpoints on a complex object”

Abstract

Aphasia is a common consequence of damage to the side of the brain, that severely limits the individual's ability to communicate by means of language. In the US (population: 311 million; source: Census bureau), approximately 80,000 adults acquire aphasia each year, and about 1 million adults currently have aphasia (source: NIDCD). By 2020, in the US (projected population: 335 million) the incidence of aphasia is expected to rise to 180,000 cases, and the prevalence to 2 million persons. This is because: population is increasing and aging; better procedures for acute neurological conditions increase survival rates; and, more effective medications and maintenance regimens increase survivors' life expectancy.

These facts create an increasing need for intervention at all stages of the disease – early diagnosis, monitoring and treatment of language deficits; development of compensatory strategies and tools – at a time when health care systems worldwide (no matter whether private or socialized) deal with a severe economic crisis. I will discuss some recent applications that may make work on disorders of speech perception and speech production more effective. I will also discuss critical aspects of the communication disabilities observed in aphasic speakers, that make research in this area different from analogous work on unimpaired communication, and that relate to a paramount need to develop simulations of human language that are both functional and biologically viable.

Presenter

Prof. Gabriele Miceli received his Medical degree from the Catholic University of Rome, and is board-certified in Neurology and in Psychiatry. He is currently professor of Neurology at the University of Trento, and clinical director of the Center for Neurocognitive Rehabilitation of CIMeC (Center for Mind/Brain Studies) in Rovereto. He had previously directed the Neuropsychological Evaluation Unit at the Institute of Neurology of the Catholic University (Rome) as an Associate professor of Neurology. Current research collaborations involve Harvard University, Johns Hopkins University, University of California at Irvine, Fondazione Bruno Kessler.

His main research area is the functional and neural organization of language functions. In the past few years, his interests have extended to the neuroimaging investigation of the neuroplasticity phenomena associated with training-induced recovery from language disorders. Present interests include the use of machine-learning algorithms in the diagnosis of language disorders, and the development of computer-assisted tools, based on conceptual lexicons, for the individualized treatment of language disorders.

Björn Granström
KTH - Royal Institute of Technology
TMH - Department of Speech, Music and Hearing, Stockholm, Sweden

“Speech technology in (re)habilitation of persons with communication disabilities”

Abstract

Speech communication research and speech technology has found many applications for handicapped individuals. One of the very first examples of an application of speech synthesis was the reading machine for the blind. It is natural that results and devices in the speech communication field can be utilized for (re)habilitation of persons with communication disabilities. AAC - Augmentative and Alternative Communication – has evolved into an independent research area with strong input from speech and language processing. In this presentation we will look at the development of the field from very early speech training devices based on speech analysis to advanced systems including robotics and avatars capable of human like interaction. We will show examples where pressing needs of disabled persons have inspired avant-garde applications and development that have eventually spread to more general use in widely used applications. In this sense the “design for all” paradigm has been a rewarding and fruitful driving force for many speech communication and technology researchers.

Presenter

Prof. Björn Granström joined the department in 1969, after graduating as MSc in Electrical Engineering. After further studies in Phonetics and General Linguistics at Stockholm University he became Doctor of Science at KTH in 1977 with the thesis "Perception and Synthesis of Speech". In 1987 he replaced Gunnar Fant as Professor in Speech Communication. He has been the director of CTT, The Center for Speech Technology, since its start in 1996. Together with Rolf Carlson, he created the first multilingual text-to-speech system, with extensive use in the disability area. Present interests include multi-modal verbal/non-verbal communication, virtual language tutors and human-like spoken dialogue systems.

Hiroshi Ishiguro
Osaka University - Dept. of Adaptive Machine Systems
Intelligent Robotics Laboratory
Dept. of Systems Innovation
Graduate School of Engineering Science
Visiting Group Leader
Dept. of Communication Robots
ATR Intelligent Robotics and Communication Laboratories

“From teleoperated androids to cellphones as surrogates”

Abstract

In order to understand the meaning of human presence, we have developed Geminoid which is an teleoperated android of myself. With the android, we could learn how people can adapt the new media. Based on the knowledge, we have recently developed a simpler teleoperated android with the minimal humanlike appearance. The new android is called Telenoid. People can easily adapt to Telenoid and enjoy conversations by using it. Further, we are remaking it with a cell-phone size. It is called Elfoid. We believe the new type of cell-phone can transmit our presence to distant places and changes our life style again.

Presenter

Prof. Hiroshi Ishiguro (M') received a D.Eng. in systems engineering from the Osaka University, Japan in 1991. He is currently Professor in the Graduate School of Engineering Science at Osaka University (2009-). He is also Group Leader (2011-) of Hiroshi Ishiguro Laboratory at the Advanced Telecommunications Research Institute, where he previously worked as Visiting Researcher (1999-2010). He was previously Research Associate (1992-1994) in the Graduate School of Engineering Science at Osaka University and Associate Professor (1998-2000) in the Department of Social Informatics at Kyoto University. He was also Visiting Scholar (1998-1999) at the University of California, San Diego, USA. He then became Associate Professor (2000-2001) and Professor (2001-2002) in the Department of Computer and Communication Sciences at Wakayama University. Then, he moved back to Osaka University as Professor in the Graduate School of Engineering (2002).

His Research Areas are sensor networks, humanoid robots, and android robots. In the last several years, he has mainly focused on development of humanoids and androids. The most successful robot he has made was exhibited in Nagoya World Expo 05, for which he is recognized as the father of the world first android. The activities have been introduced by almost all major TVs, such as BBC, and newspapers in the world. Recently, he has been selected as one of 100 geniuses alive in this world by Synectics (UK, 2007).