10. Interspeech 2008 Special Sessions

Rooms: BCEC PLAZA 3 & 4

Phonetics and Phonology of Australian Aboriginal languages (OzPhon08)

TueSe3.O4 Tuesday 13:30-15:30

The aim of this special session is to foster collaboration between those concerned with fundamental mechanisms of speech production and perception and those who are primarily concerned with the description of sound systems of indigenous Australian languages. Recent research has suggested that the sound systems of many indigenous languages of Australia are sufficiently different from those of the better-studied European and Asian languages to raise the prospect that the former impose unique demands on listeners, with implications for models of speech processing. At the same time, there is also a lack of experimental data and an urgent need to document the phonetic and phonological properties of indigenous Australian languages, in the face of massive and ongoing language attrition.

Organiser

Marija Tabain, La Trobe University

The Phonetics and Phonology of Australian and New Zealand English (PANZE 2008)

TueSe4.O4 Tuesday 16:00-1800

The aims of the workshop are to bring together researchers who use experimental methods in the examination of Australian and/or New Zealand English, to discuss techniques and procedures, and to exchange findings on Australian and New Zealand English with respect to any of the following topics:

- Sound Change
- Intonation
- Regional Variation
- Sociocultural Variation
- Comparative Studies

Organisers

Felicity Cox, Macquarie University
Sallyanne Palethorpe, Macquarie University
Catherine Watson, University of Auckland

Cross-linguistic and developmental issues in the perception and production of lexical tone

WedSe2.O4 Wednesday 10:00-12:00

Perception of consonants and vowels is well studied in psycholinguistic, neurolinguistic, and developmental contexts, but by comparison, tone is relatively understudied. The study of tone processing is important first because there is a void in our knowledge of how the majority of the world's population come to be fluent perceivers and producers of tone in their native language, and second, understanding how the linguistic environment influences native, non-native, and L2 tone processing (perception and production) will allow descriptions of speech perception, language development, and the nature of psycholinguistic processes in human cognition to be expanded.

Organiser

Karen Mattock, Lancaster University

Auditory-inspired spectro-temporal features

WedSe3.04 Wednesday 13:30-15:30

Psychoacoustic models had a strong influence on the development of speech features in the past. Features inspired by neurophysiologic findings, namely spectro-temporal features, capturing the spectral and dynamical aspects of speech, have the potential to be of similar influence in the future. This special session is dedicated to the neural basis of spectro-temporal features and their use in speech processing. Following the plenary talk by Sophie Scott at INTERSPEECH 2007 and the panel discussion by Anne Cutler and Roger Moore at INTERSPEECH 2005, we hope this session will further vitalize the exchange between traditional participants of INTERSPEECH and people interested in the neuroscience of speech.

Organiser

Martin Heckmann, Honda Research Institute Europe GmbH

The Prosody of Spontaneous Speech

WedSe4.O4 Wednesday 16:00-18:00

Traditional quantitative prosodic models have typically been developed on text-elicited "lab speech" whose prosody is rather closely linked to linguistic structures often predefined by lexical constituency and punctuation. In contrast, spontaneous speech, with its multi-tiered, rich information structure has mostly been analysed, at best, impressionistically. Recently, however, the analysis and synthesis and recognition of more varied as well as affective speaking styles both of which are, in a completely natural way, connected with spontaneous utterances, has drawn growing attention. This means that also models of e.g. f0, speech rate, as well as voice quality have to transcend the realm of traditional linguistics and incorporate discourse-related as well as para-linguistic information.

Organisers

Hansjörg Mixdorff, TFH Berlin University of Applied Sciences, Germany
Hartmut Pfitzinger, University of Kiel, Germany
Keikichi Hirose, University of Tokyo, Japan

Forensic Speaker Recognition - Traditional and Automatic Approaches

ThuSe2.O4 Thursday 10:00-12:00

Two basic approaches to forensic speaker recognition are a) the traditional, rooted in phonetics, and b) the automatic, rooted in engineering. This special session will include papers from both approaches, and will consider how the strengths of the two approaches may be combined.

Organiser

Geoffrey Stewart Morrison, School of Language Studies, Australian National University

Panel Discussants

Joaquín González-Rodrígez
Biometric Recognition Group - ATVS, Universidad Autónoma de Madrid
http://arantxa.ii.uam.es/~joaquin/contactar.html

Takashi Osanai
National Research Institute of Police Science, Japan
http://www.nrips.go.jp/

Philip Rose
School of Language Studies, Australian National University
http://arts.anu.edu.au/languages/linguistics/Phil_Rose.asp

Elizabeth Shriberg
Speech Technology & Research International / International Computer Science Institute
http://www.speech.sri.com/people/ees/

Human-Machine Comparisons of Consonant Recognition in Noise.

ThuSe3.O4 Thursday 13:30-15:30

Listeners outperform automatic speech recognition systems at every level of speech recognition, including the very basic level of consonant recognition. What is not clear is where the human advantage originates. The purpose of this Special Session is to promote focused human-computer comparisons on a task involving consonant identification in noise, with all participants using the same training and test data. Training and test data and native listener and baseline recogniser results will be provided by the organisers, but participants are encouraged to also contribute listener responses.

Organisers:

Odette Scharenborg, Radboud University Nijmegen
Martin Cooke, University of Sheffield

LIPS 2008: Visual Speech Synthesis Challenge

FriSe2.O4 Friday 10:00-12:00

LIPS 2008: the first visual speech synthesis challenge will be held as a special session at INTERSPEECH 2008 in Brisbane, Australia. The aim of this challenge is to stimulate discussion about subjective quality assessment of synthesised visual speech with a view to developing standardised evaluation procedures.

Organisers:

Sascha Fagel, Berlin University of Technology
Barry-John Theobald, University of East Anglia
Gérard Bailly, GIPSA-Lab, Dept. Parole et Cognition, Université Stendhal Grenoble 3

Talking heads and pronunciation training

FriSe3.O4 Friday 13:30-15:30

When first experiencing Abbé Mical's talking heads in 1783, the journalist Rivarol envisioned a future where such mechanical speech production will be set in all houses and will teach French people and foreigners a standardized language with no dialectal variations and no passions! If TV sets are quite addressing most of the challenge, talking heads can be quite helpful tools in language training and clinical phonetics. The objective of this special session is to have an overview of the potential uses of embodied speech synthesis systems as virtual language tutors. The special session is mainly focused on experimental results measuring the impact of augmented multimodal speech displays on language training, from second language learning (including cued speech or sign languages) to phonetic correction.

Organiser

Gérard Bailly, GIPSA-Lab, Dept. Parole et Cognition, Université Stendhal Grenoble 3