SPECOM 2026 Program at a Glance

17 - 18 September 2026, Ohrid, North Macedonia

Thursday 17 September 2026

08:00 - 09:00	Registration	
09:00 - 09:15	**Opening ceremony**	
09:15 - 10:15	**Keynote talk** *Artificial Intelligence versus Human Intelligence, from the Ground up* **Yi Xu**, University College London, London, UK	
10:15 - 11:45	**ASR**	**Vocal Biomarkers 1**
11:45 - 12:00	Coffee break	
12:00 - 13:30	**Phonetic Analysis**	**Vocal Biomarkers 2**
13:30 - 14:30	Lunch break	
14:30 - 16:00	**Speech Signal Processing**	**LLM**
17:00 -	**Boat trip** and **Conference Dinner**	

Friday 18 September 2026

08:30 - 10:00	**NLP**	
10:00 - 11:00	**Keynote talk** *From Prosody to Spiking Networks: Biological Plausibility in Speech and Audio Processing* **Phil Garner**, Idiap Research Institute, Martigny, Switzerland	
11:00 - 11:15	Coffee break	
12:15 - 13:45	**TTS**	**Emotion Recognition**
13:45 - 14:45	Lunch break	
14:45 - 16:15	**Multimodal Processing**	**Speaker ID and Resources**
16:15 - 16:30	**Closing ceremony**	

SPECOM 2026 Detailed Session Program

Thursday 17 September 2026

10:15 - 11:45 **ASR** (Conference Hall 1)

Chairs: Tibor Fegyó and Andrej Zgank

How Robust are Modern Automatic Speech Recognizers to Domain and Temporal Mismatch? A Study on the Less-resourced Slovenian Language
Andrej Zgank, Urh Kolaric and Zdravko Kacic

LoRA-Based Speech-Augmented Language Model Adaptation for Hungarian ASR
Mate Kadar, Gergely Dobsinszki, Tibor Fegyo and Peter Mihajlik

Sequential Domain Adaptation of Whisper for Kazakh Academic Lecture Speech Recognition
Zhanibek Kozhirbayev and Zhandos Yessenbayev

A System for Automatic Transcription of Macedonian Parliamentary Speech
Ema Mishevksa, Marija Markovska Dimitrovska, Dimitar Taskovski, Tanja Danailovska and Branislav Gerazov

Understanding Personalization Gains in Whisper ASR
Sergei Katkov and Antonio Liotta

Cross-Lingual Robustness of Zipformer: Czech and German ASR Benchmarking Study
Jan Lehečka, Lubos Smidl, Josef V. Psutka and Josef Psutka

10:15 - 11:45 **Vocal Biomarkers 1** (Conference Hall 2)

Chairs: Alexey Karpov and Mikhail Dolgushin

Explainable Cognitive Impairments Classification in Russian Speech
Mikhail Dolgushin, Irina Kipyatkova, Alexey Karpov, Dmitry Boldakov and Igor Dorovskikh

PA-Profile: Accessible, Screener-Agnostic Speech-Sound-Disorder Screening with ASR-Derived Features in English and Low-Resource Kannada
Deepak Swaroop, Veena Thenkanidiyoor, Dileep A D and Muralikrishna HV

eGeMAPS-Based Detection of Alzheimer's Disease in Russian-Language Speech: A Pilot Study
Elena Nikolaeva, Karina Evgrafova, Vera Evdokimova, Iurii Poliakov and Polina Soroka

Exploring Voice and Speech as ADHD Markers
Elena Lyakso, Olga Frolova, Anton Matveev, Egor Kleshnev, Aleksandr Nikolaev, Platon Grave, Abylay Ilyas and Olesia Koroteeva

Electroglottographic Features of Children with ADHD: a Pilot Study
Abylay Ilyas, Olga Frolova, Anton Matveev and Elena Lyakso

ASR Personalization in Pathological Voice: Accuracy, Efficiency, and Duration Effects
Giuseppe Giaimis, Sergei Katkov, Muhammad Azfar Yaqub and Antonio Liotta

12:00 - 13:30 **Phonetic Analysis** (Conference Hall 1)

Chairs: Anna Oleynik and Anna Borzykh

Rhythmic Accommodation in English as a Lingua Franca: A Case Study of German–Italian Dialogue
Anna Oleynik and Anna Borzykh

Predicting changes in L1 vowels of late Czech and Italian bilinguals with French as L2
Marie Fongaro, Barbara Gili Fivela, Maud Pélissier and Gabriel Hévr

Pilot Classification of Macedonian Regional Speech based on Acoustic Features and Machine Learning
Elena Todorova, Cvetanka Bilbiloska, Ema Mishevska, Veselinka Labroska, Tomislav Kartalov, Katerina Raleva, Dimitar Taskovski and Branislav Gerazov

Chinese Speakers' Adaptability to English Accents: a Perceptual Analysis of Word Stress Patterns
Daria Artyukhova and Tatiana Shevchenko

Rhythm Contributes to Personality Identification: Gender and Age Features in Australian and New Zealand Corpora
Tatiana Shevchenko and Anna Borzykh

Acoustic Distinctiveness of Dialects: A Machine–Learning Approach to Clustering Southern Min Varieties
Man-Ni Chu, Chang-Han Lin and Jia-Lien Hsu

12:00 - 13:30 **Vocal Biomarkers 2** (Conference Hall 2)

Chairs: Wolfgang Minker and Evgeny Kostyuchenko

Effects of Dysarthria on Speech Prosody - A Case Study
Laszlo Toth, Gábor Gosztolya and Andrea Deme

Personalized Speech Recovery Tracking with Wav2vec 2.0: A Longitudinal Study on Head-and-Neck Surgery Patients
Evgeny Kostyuchenko

Prosodic Features for Classification of Normal vs Dysarthric Speech
Manya Singh, Avantika Agarwal and Hemant Patil

VocalPlan-AV: A Two-Stage Audio-Visual Framework for Dysarthric ASR Using Vocal-Plan Reconstruction
Dev Jani, Dhaneshkumar Kapadia, Samarth Modh, Nisarg Trivedi and Hemant A Patil

Cross-Lingual Dysarthria Detection: Rigorous Evaluation Methodology and Transfer Analysis Across Four Languages
Ivan Gusev, Anastasia Pivovarova and Yuri Matveev

Impact of Audio Codecs on Automatic Depression Detection: A Cross-Corpus Study
Danila Mamontov, Dennis Gaebler, Alexey Karpov and Wolfgang Minker

14:30 - 16:00 **Speech Signal Processing** (Conference Hall 1)

Chairs: Geza Nemeth and Denis Ivanko

Robust Speech Enhancement via Extended LSTM with Perceptual Contrast Stretching and Monte Carlo Dropout
Denis Ivanko and Minglei Fu

Multiple Noise Suppression via Spatiotemporal Spectral Image Inpainting
Yuma Morozumi, Teruki Toya and Kenji Ozawa

Restoring Bone-Conducted Whispered Speech Using Self-Supervised Voice Conversion with Pseudo Bone-Conduction Data Augmentation
Akihiko Kusada, Kenji Ozawa and Teruki Toya

A Lightweight Voice Activity Detection Model for Scene-Aware Cochlear Implant Processing
Kashif Ali, Garima Verma, Geza Nemeth and Paul Oomen

Applicability and Generalization of Audio Fingerprinting Methods to Speech
Kemal Altwlkany, Elmedin Selmanovic, Sead Delalic and Emanuel Lacic

14:30 - 16:00 **LLM** (Conference Hall 2)
Chairs: Ivan Kraljevski and Milan Sečujski

Iterative LLM-based improvement for French Clinical Interview Transcription and Speaker Diarization
Ambre Marie, Thomas Bertin, Guillaume Dardenne and Gwenolé Quellec

Exploiting large language models in safety-critical medical helplines: a study in Serbian and English
Milan Sečujski, Aleksandar Vujinović, Aleksandar Kovačević, Nikola Simić, Vuk Stanojev, Siniša Suzić, Tijana Nosek and Vlado Delić

Utilizing LLMs for Domain Specific Corpus Augmentation for Sorbian Languages
Ivan Kraljevski, Talib Gasanov, Daniel Sobe, Frank Duckhorn, Constanze Tschöpe and Matthias Wolff

AI vs. Human Emotional Narratives: A Comparative Study of Subclausal Units in LLM-Generated and Human Spoken Texts
Anna Leonteva, Veronika Izumskaya-Kapitonova and Maria Kusina

Comparing Large Language Models for Aspect-Based Sentiment Analysis of Student Experience from Russian-Language Online Course Reviews
Varvara Kalicheva and Margarita Kirina

UzEnParC 1.0: A Cross-Script, Metadata-Rich Uzbek-English Parallel Corpus for Low-Resource NLP
Marufjon Amirkulov and Botir Elov

Friday 18 September 2026

08:30 - 10:00 **NLP** (Conference Hall 2)

Chairs: Yannick Estève and Salima Mdhaffar

A Multi-Stage Method for Training Semantic Search Models
Sergei Leshchenko, Sergei Masliukhin, Olesia Koroteeva and Yuri Matveev

LEMON: Factor-Level Semantic Alignment for Graph-Text Pairs
Anton Tomilov, Daria Gineva, Danil Tirskikh, Olesia Koroteeva and Yuri Matveev

Hybrid Algorithm for Automatic Annotation of Modified Multiword Expressions in Russian Colloquial Speech (Based on the ORD Corpus)
Elizaveta Zakharina and Tatiana Popova

Synchronic and Diachronic Models of Motivation Ratings of English Words
Anna Shevlyakova, Vladimir Bochkarev and Andrey Savinkov

Semantic Selection and Filtering of Pseudo-Labelled Data for Low-Resource Speech Translation
Othman Istaiteh, Salima Mdhaffar and Yannick Estève

Computer-Aided Language Learning for Arabic-Speaking Children
Jouwana Daibes, Klarien Wassaya, Masa Jalamneh and Abualsoud Hanani

12:15 - 13:45 **TTS** (Conference Hall 1)

Chairs: Alexey Karpov and Yunus Emre Özköse

A Comprehensive Objective Evaluation of Modern Text-to-Speech for Turkish Using Speech Quality Assessment Models
Yunus Emre Özköse, Alperen Kahraman and Ali Haznedaroglu

Towards Streaming Neural Speech Codecs through Time-Invariant Representations
Kélian Estève, Salima Mdhaffar, Mickael Rouvier, Richard Dufour and Yannick Estève

AccentFlow: Continuous Accent Transfer via Accent Delta Vectors in Speaker Embedding Space
Artem Ostrovskii, Arseniy Sheremet and Alexey Karpov

Multi-Style Child TTS: An Expressive Text-to-Speech System for Children in Low-Resource Scenarios
Shaimaa Alwaisi, Mohammed Salah Al-Radhi and Géza Németh

A Lightweight LSTM-Based iEEG-to-Speech Decoder with Event-Level Unseen Neural Pattern Detection
Saifullah Memon and Mohammed Salah Al Radhi

The Neural Signature of Dark Humor Memes: Beta Oscillations in the Prefrontal Cortex and Temporoparietal Junction
Vsevolod Potapov, Ekaterina Karimova, Rodmonga Potapova, Diana Smolskaya, Nikolay Bobrov, Leonid Motovskikh and Iurii Pozhilov

12:15 - 13:45 **Emotion Recognition** (Conference Hall 2)

Chairs: Gábor Gosztolya and Anton Yakovenko

Detecting Sarcasm From Speech Using Transformer-Based and Traditional Methods
Gábor Gosztolya, Márton Makrai, Mercedes Kiss-Vetráb and László Balázs

Automatic Classification vs. Human Annotation of Emotions in Everyday Spoken Russian: A Case Study of the ESC Corpus
Margarita A. Kirina and Sofiia Chepovetskaia

Effect of the Amount of Training Data in Speech Emotion Recognition
Dániel Halmai and Gábor Gosztolya

Emotion-Trajectory-Driven Segment Selection for Depression Severity Prediction
Jiawei Yu and Heysem Kaya

AMYGDALA: A Speech Foundation Model Adaptation Framework for Robust Emotion Recognition in Low-Resource Scenarios
Anton Yakovenko and Evgeniy Shuranov

An Audio-Gated Cascade for Resource-Efficient Multimodal Emotion Recognition
Damir Kandelov and Lyudmila Savchenko

14:45 - 16:15 **Multimodal Processing** (Conference Hall 1)

Chairs: Vlado Delić and Andrea Corradini

A Unified Phoneme-to-Viseme Mapping for Serbian Based on Linguistic Knowledge, Human Perception, and AV-HuBERT Visual Embeddings
Siniša Suzić, Nosek Tijana, Nikola Simić, Vuk Stanojev, Nikša Jakovljević, Vlado Delić and Milan Sečujski

Talking Faces in Serbian: From Speech and Text to Blendshape Animation
Nikola Stojičić, Milica Bošnjak, Tijana Petrović, Iva Jovanović, Siniša Suzić and Tijana Nosek

The Influence of Implicit Robotic Cues on the Efficiency of Cognitive Task Performance
Ekaterina Dubinina and Artemiy Kotov

Understanding Pose-Based Sign Language Translation: Ablation Studies
Tomáš Železný, Jakub Straka, Václav Javorek, Ondřej Valach, Marek Hrúz and Ivan Gruber

Gaze, Speech and Gesture of L2 Learners as Mediated by a Discourse Type: Description, Narration and Instruction
Maria Kiose, Anastasia Rzheshevskaya and Vadim Potekhin

Sound and Colour: Acoustic and Visual Features of Vowel Perception in Mongolian in Comparison with Russian
Rodmonga Potapova, Vsevolod Potapov, Tsend-Ayush Ganbaatar and Leonid Motovskikh

14:45 - 16:15 **Speaker ID and Resources** (Conference Hall 2)

Chairs: Peter Mihajlik and Piroska Zsófia Barta

When Pseudo-Labels Go Wrong: The Importance of Label Correction in Self-Supervised Speaker Verification
Abderrahim Fathan and Jahangir Alam

Investigating Transformers-based Feature Representations for Audio-Visual Speaker Verification
Faqrul Alam Chowdhury and Jahangir Alam

Multilingual Voicemail and Answering Machine Detection via Pretrained Audio Embeddings and Gradient Boosting Machines
Kemal Altwlkany and Muamer Paric

StreamSV: End-to-End Streaming Speaker Verification for Short Utterances
Stepan Malykh and Sergey Novoselov

Scaling Conversational Hungarian ASR: The BEA-Dialogue+ Corpus
Máté Gedeon, Piroska Zsófia Barta, Péter Mihajlik and Katalin Mády

A Corpus-Based Approach to Candidate Generation for Korean Dialect ASR Post-Processing
Eungseok Kim, Sejeong Hong, Sangmi Lee and Su Cha