

IMPACT-SPEECH 2026

**IMPACT-SPEECH: Identifying, Measuring, Preventing, and
Assessing Consequences of Bias in Speech LLMs**

Proceedings of the Workshop

October 29, 2026

©2026 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-479-8

Preface

This is the first edition of the IMPACT-SPEECH Workshop, co-located with the 2026 Conference on Empirical Methods in Natural Language Processing (EMNLP 2026) and held on October 29, 2026, in Budapest, Hungary.

We are excited to welcome attendees to the IMPACT-SPEECH 2026 Workshop on Identifying, Measuring, Preventing, and Assessing Consequences of Bias in Speech LLMs. The workshop brings together researchers and practitioners working across speech processing, natural language processing, machine learning, fairness, and responsible AI to examine the growing challenges associated with bias in speech-enabled large language models and multimodal speech technologies.

As speech AI systems become increasingly integrated into real-world applications, they interact with speakers from diverse linguistic, demographic, cultural, and social backgrounds. These interactions raise important questions about how bias in such systems can be identified and measured, how developers can prevent or mitigate it, and how biased system behavior can affect individuals and communities. IMPACT-SPEECH 2026 provides a forum for researchers to discuss these questions and share approaches for building fairer, more inclusive, and reliable speech technologies.

This year’s program features 22 archival and non-archival papers, including 8 oral and 14 poster presentations, alongside 2 invited talks. Together, these contributions cover a broad range of topics related to bias and fairness in speech LLMs. We are grateful to all authors and participants for contributing their research, ideas, and perspectives to the workshop.

We also extend our sincere thanks to the EMNLP 2026 organizers for providing the opportunity to hold IMPACT-SPEECH alongside the conference and for their considerable efforts in creating an environment that supports discussion and collaboration across the NLP and speech communities.

Finally, we thank all members of the IMPACT-SPEECH 2026 Program Committee for their time, expertise, and thoughtful reviews. Their contributions have helped us curate a diverse and engaging program. We hope that this workshop will stimulate new collaborations, encourage further research, and contribute to the development of more equitable and responsible speech AI.

The IMPACT-SPEECH 2026 Organizing Committee: Ravi Shekhar, Monorama Swain, Jagabandhu Mishra, Sandipan Dhar, Haralambos Mouratidis, and Matthew Purver

Organizing Committee

Workshop Organizers

Ravi Shekhar, University of Essex, UK

Monorama Swain, JKU Linz, Austria

Jagabandhu Mishra, University of Eastern Finland, Finland

Sandipan Dhar, Fraunhofer IIS, Germany

Haralambos Mouratidis, University of Essex, UK

Matthew Purver, QMUL, UK and JSI, Slovenia

Program Committee

Program Committee

Badr M. Abdullah, Saarland University, Germany
Mohd Mujtaba Akhtar, IIIT Delhi, India
Iqra Ali, Queen Mary University of London, UK
Kalluri Shareef Babu, Manipal Institute of Technology Bengaluru
Alessio Brutti, Fondazione Bruno Kessler (FBK), Italy
Sergio Burdisso, Idiap Research Institute, Switzerland
Shammur Absar Chowdhury, Qatar Computing Research Institute (QCRI), Qatar
Lorenzo Concina, Fondazione Bruno Kessler (FBK), Italy
Sudhansu Bala Das, University of Galway, Ireland
Giovanni Filomeno, JKU Linz, Austria
Girish, IIIT Delhi, India
Reyhaneh Hashempour, QA Ltd, UK
Daniel Hershcovich, University of Copenhagen, Denmark
Nina Hosseini-Kivanani, Radio Télévision Luxembourg (RTL), Luxembourg
Parisa Jamadi, Imperial College London, UK
Vanja Karan, University of Vienna, Austria
Santosh Kesiraju, Brno University of Technology, Czechia
Sandeep Kumar, MBZUAI, UAE
Spyretta Leivaditi, University of Groningen, Netherlands
Yunfei Long, Queen Mary University of London, UK
Fernando López, Telefónica Innovación Digital, Spain
Ben Malin, Brunel University, UK
Shah Nawaz, Johannes Kepler University Linz, Austria
Shekhar Nayak, University of Groningen, Netherlands
Ludovica Pannitto, University of Bologna, Italy
Tanvina Patel, TU Delft, Netherlands
Orchid Chetia Phukan, IIIT Delhi, India
Senja Pollak, Jožef Stefan Institute, Slovenia
Priyanshu Priya, IIT Patna, India
Tariq Rahim, Kingston University London, UK
Tharindu Ranasinghe, Lancaster University, UK
Subhankar Roy, University of Bergamo, Italy
Markus Schedl, Johannes Kepler University Linz, Austria
Divya Sharma, IIIT Delhi, India
Samara Souza, Federal University of Pará, Brazil
Bo Su, Indiana University Bloomington, USA
Abhirama Subramanyam, MBZUAI Abu Dhabi
Richard Sutcliffe, University of Essex, UK
Eva Szekely, KTH Royal Institute of Technology, Sweden
Venktesh V, Stockholm, Sweden
Esaú Villatoro Tello, Idiap Research Institute, Switzerland
Sarenne Wallbridge, UK AI Security Institute, UK
Tariq Yousef, University of Southern Denmark, Denmark
Juntao Yu, Queen Mary University of London, UK
Yuchen Zhang, University of Essex, UK

Keynote Talk

Every Voice Matters: Towards Inclusive Speech LLMs

Marco Matassoni

Fondazione Bruno Kessler (FBK), Trento, Italy



Abstract: The increasing use of Speech Large Language Models raises fundamental questions about fairness, inclusiveness, and trustworthiness. This talk presents research conducted within the ELOQUENCE project, which aims to develop multilingual, context-aware, and bias-controlled conversational AI for applications ranging from customer service and smart homes to safety-critical scenarios. We will focus on inequalities in multilingual and low-resource speech processing, where differences in data availability, language coverage, accents, and speaker characteristics can lead to significant performance disparities. The talk will discuss approaches to identify and measure speech bias, as well as strategies to improve robustness and inclusiveness through cross-lingual adaptation, data-efficient learning, and privacy-preserving approaches.

Bio: Marco Matassoni is a Senior Researcher at the Fondazione Bruno Kessler (FBK) in Trento, Italy, where he plays a leading role within the Speech Technology (SpeechTeK) research unit of the Centre for Augmented Intelligence. Over nearly three decades, he has become an experienced figure in the field of speech processing, contributing to advancements in automatic speech recognition (ASR) and human-machine interaction systems. Marco Matassoni graduated in Electronic Engineering from the University of Padua in 1995, where his thesis focused on microphone arrays, a topic that would later define part of his professional research. In 1996, he joined ITC-IRST, the research institute that later became part of FBK, and began working on speech recognition technologies. His research was driven by the challenge of making speech systems perform robustly in real-world environments marked by noise, reverberation, and distance between speakers and microphones. During his career, he has contributed to a wide range of European research projects, including SpeechDat-Car, VICO, HIWIRE, DICIT, DIRHA. These initiatives focused on developing speech interfaces for complex acoustic conditions such as in-car communication or home automation. His work on distant-talking and far-field speech recognition has been particularly influential, especially on data augmentation techniques. In recent years, he has expanded his focus to include multilingual and low-resource ASR, speaker verification and anonymization, and bias-aware speech technologies; in the ELOQUENCE project, he is working on multilingual and cross-lingual speech recognition, addressing the challenges of building scalable models for European languages. These efforts reflect a growing concern with inclusive artificial intelligence, ensuring that speech technologies respect linguistic diversity and privacy. In addition, he regularly serves as a reviewer for major conferences and journals in the field, contributing his expertise to the peer-review process and supporting the dissemination of high-quality research.

Keynote Talk

Bias before the model: Reference noise as source of unfair speech systems

Jesper Rindom Jensen
Aalborg University



Abstract: Bias in speech systems is often attributed to the models themselves, but it can be introduced long before training begins. In this talk, I will use speech enhancement as a case study, focusing on one such example: the clean references used for training are, in practice, often noisy. Drawing on our recent research, we will see how this can bias models to let noise pass through while still scoring well on objective metrics. We will also examine current evaluation practices, showing that common metrics struggle to differentiate between state-of-the-art methods and often fail to reflect the preferences and perception of real users.

Bio: Jesper Rindom Jensen was born in Ringkøbing, Denmark in August 1984. He received the M.Sc. degree cum laude in 2009 and the Ph.D. degree in 2012, both from Aalborg University, Denmark. He is currently an Associate Professor at the Department of Electronic Systems at Aalborg University, where he is enrolled in the Professor Promotion Programme and serves as Head of the Doctoral Programme in Electrical and Electronic Engineering at the Technical Doctoral School of IT and Design. He is also Head of Research at the Audio Analysis Lab, which he has been a member of since its foundation in 2012. Before this, he was with the Department of Architecture, Design & Media Technology at Aalborg University as a Postdoc, Assistant Professor, and Associate Professor. He has been a Visiting Researcher at the University of Quebec, INRS-EMT, in Montreal, Quebec, Canada, at the Friedrich-Alexander Universität Erlangen-Nürnberg in Erlangen, Germany, and at the University of Surrey, UK. His research interests include speech and audio processing, multichannel and array processing, robot and drone audition, spatial audio, estimation theory, noise reduction, and machine learning for signal processing. He has published more than 120 papers in top-tier, peer-reviewed conference proceedings and journals, holds a patent, and is the co-author of two books, namely, *Speech Enhancement – A Signal Subspace Perspective* and *Signal Enhancement with Variable Span Linear Filters*. He has led a number of externally funded research projects, including as Principal Investigator and Coordinator of a European Defence Fund project on AI-based natural language processing. Dr. Jensen has received a highly competitive individual postdoc grant from the Independent Research Fund Denmark, has been selected as an AAU Talent at Aalborg University (awarded to young research talents), was named Teacher of the Year by the Study Board of Media Technology, and has received several travel grants from private foundations. He is a member of the IEEE Speech and Language Processing Technical Committee and of the EURASIP Technical Area Committee on Acoustic, Speech, and Music Signal Processing, and is a Senior Member of the IEEE. He currently served as an Associate Editor of the Elsevier Signal Processing Open journal and has previously served as a member of the IEEE Audio and Acoustic Signal Processing Technical Committee, and as an Affiliate Member of the IEEE Signal Processing Theory and Methods Technical Committee, and he has served as area or track chair for ICASSP, WASPAA, and EUSIPCO.

Table of Contents

<i>A Probe Shift Is Not a Fairness Fix: The Limits of Representation Steering in Speech Models</i> Nicolas Bourrel, Abderrahmane Issam and Gerasimos Spanakis	1
<i>AIP-Speech: Auditing Speech LLM Robustness to Background Interference</i> Aiden Dsouza, Shreyansh Mishra, Yashraj Suneja, Rohan Swain, Samyak Borkar, Arjun Jaishankar and Tanmay Urade	10
<i>Can Large Audio-Language Models Ignore Multilingual Distractors? An Evaluation of Their Selective Auditory Attention Capabilities</i> Heejoon Koo	20
<i>Early Warning for Target-Speaker Extraction: Reliability under Distribution and Failure-Definition Shift</i> Shubham Sharma and Manik Gupta	33
<i>Evaluating Bias in Phoneme-Based Automatic Speech Recognition Systems: An Analysis of IPA Transcription Models</i> Maneesha Rani Saha, Catherine Bao and Neal Patwari	42
<i>Lost in Compression: Calibration Bias and the Accent-WER Gap in Speech LLMs</i> Angad Ripudaman Singh Bajwa	61
<i>Pruning for Efficiency, Paying in Fairness: Demographic Disparities in Pruned Speech-LLMs</i> Ganesh Pavan Kartikeya Bharadwaj Kolluri, Michael Kampouridis and Ravi Shekhar	67
<i>Temporal Taxation Compounds Under Post-Training Compression of Whisper Models</i> Srishti Ginjala, Eric Fosler-Lussier, Christopher W. Myers and Srinivasan Parthasarathy	81
<i>TurkSV-Bench: Assessing Demographic Bias and Preprocessing Effects in Speaker Verification</i> Ahmet Berkay Kantarcı, Abdulkadir Parlak and Ahmet Karazor	95
<i>What Drives Dialectal Jailbreaks? An Ablation of Surface Form, Cultural Framing, and Strategy Banks</i> Qingyang Xu	104