

**IWSDS 2026**

**The 16<sup>th</sup> International Workshop on  
Spoken Dialogue Systems Technology**

**Proceedings of the Conference**

February 26 - March 1, 2026  
Trento, Italy

©2026 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 978-1-952148-25-5

## Message from the General Chair

The 16<sup>th</sup> International Workshop on Spoken Dialogue Systems Technology (IWSDS 2026) was held on February 26 - March 1, 2026 in Trento, Italy. This year's conference theme was "**Human-Machine Dialogue in the Era of Multimodal Foundation Models**". IWSDS 2026 benefited from the experience and knowledge generated by previous editions since 2009.

The IWSDS conference series is a consolidated forum where international researchers, practitioners, and stakeholders working in the field of spoken dialogue systems and associated technologies can disseminate their current research and applications, discuss technological challenges, present their success stories, and share their complementary visions for the technology's future. Human-Machine Dialogue will be the key enabler of many technologies and tools developed to support the expectations of AI transformation. Whether it is a dialogue in the medicine, health, robotics, or virtual worlds domain, we need to rethink core computational and interaction models. IWSDS 2026 has been fertile ground for meetings, for elaborating on challenges, and for grounding a shared vision that benefits humanity.

The conference invited and received submissions grouped into four categories: a) long research papers, reporting mature and completed research results; b) short research papers, presenting focused studies or ongoing but original research efforts; c) position papers, describing novel research ideas or viewpoints intended to stimulate discussion; and d) demonstration papers, presenting implemented systems or prototypes. These contributions addressed a broad range of topics related to the theoretical foundations, systems and methods, and applications of spoken and multimodal dialogue systems. Topics of interest included, but were not limited to, the following areas::

- **Theoretical Foundations:**

- Multimodal and situated dialogue models
- Integration and control of modalities and context
- Incremental and longitudinal dialogue models
- Emotion, engagement, and interaction styles
- Ethical, trustworthy, and responsible dialogue systems
- Bias, toxicity, and harm in dialogue systems

- **Systems and Methods:**

- Multimodal Models (foundation, small or task-specific pre-trained models) for Dialogue Systems
- Multimodal Models for grounding, personalization, safety, and explainability
- Dialogue generation and management
- Grounded and adaptive Natural Language Generation
- Domain transfer and adaptation techniques

- Evaluation methods and multimodal metrics
- Resources, corpora creation, annotation
- **Applications:**
  - Educational and healthcare applications
  - Low-resource and multilingual settings
  - Robotics and embodied dialogue systems
  - Dialogue in Virtual and Extended Reality
  - Real-world dialogue systems

In addition to the regular research tracks, the conference featured a dedicated **Industry Track** to present practical experiences with spoken and multimodal dialogue systems in real-world settings. Contributions to this track focused on deployed systems, prototypes, integration efforts, and user studies, highlighting challenges encountered during development and deployment, as well as lessons learned in industrial or commercial contexts. Submissions to the Industry Track were not required to release data or source code; however, clear descriptions of the systems, experimental settings, and evaluation results were expected. Industry Track papers were reviewed primarily for practical relevance, robustness, innovation, and impact, rather than solely for methodological novelty.

IWSDS 2026 received a total of **55 submissions**. Each submission was reviewed by at least three members of the Program Committee. Based on the review process, 42 papers were accepted for presentation at the workshop. The accepted contributions were distributed as follows: 15 papers were presented in the three General Track sessions (*Dialogue and Interactive Systems*, *Resources and Evaluation*, and *Human-Centered Interaction*); 12 papers were presented in three Special Sessions (*Human-Machine Dialogue in the Era of Multimodal Foundation Models*, *Conversational AI Applications in Robotics & Virtual Reality*, and *Conversational AI Applications in the Health Domain*); 12 papers were presented in poster sessions; and 3 papers were accepted for the Industry Track.

The workshop recognized outstanding contributions through a "Best Paper Award" and an "Outstanding Paper Award". The awards were selected based on reviewers' evaluations and the Program Committee's assessment, with consideration for originality, technical quality, and relevance to the workshop themes.

The workshop featured **three keynote talks** addressing complementary perspectives on spoken and multimodal dialogue systems: a) *"Toward Socially Intelligent Dialogue Systems: From Foundational Models to Mental Models"*, by Asli Celikyilmaz (Meta Fundamentals AI Research); b) *"Towards More Human-like Conversational Agents: Opportunities and Challenges"*, by Gabriel Skantze (KTH Royal Institute of Technology); and, c) *"Towards Meaningful Control of Multimodal XR Experiences and the Role of Spoken Dialogue"*, by Giulio Jacucci (University of Helsinki).

In addition to the regular paper sessions, IWSDS 2026 included **three Special Sessions** addressing emerging and application-driven directions in spoken and multimodal dialogue systems. The Special Session *"Human-Machine Dialogue in the Era of Multimodal Foundation Models"* focused on the opportunities and challenges posed by large-scale multimodal models, including issues related to grounding, interaction dynamics, evaluation, and robustness in multi-turn, multi-party dialogue. The Special Session *"Conversational AI Applications in Robotics & Virtual Reality"* addressed conversational AI in embodied and immersive environments, with contributions covering human-robot interaction, adaptive dialogue in physical and virtual spaces, and multimodal perception and action. The Special Session *"Conversational AI Applications in the Health Domain"* focused on dialogue systems for healthcare applications, including clinical decision support, patient-facing conversational agents, and ethical and safety considerations in health-related dialogue technologies.

Complementing the Special Sessions, the program featured **three panel discussions** corresponding to the same themes. These panels provided a forum for discussion among researchers and practitioners from academia and industry, addressing open research challenges, methodological issues, and future directions for spoken and multimodal dialogue systems in each domain.

The organizers would like to thank the chairs of the IWSDS26 Committee for their support and advice throughout the preparation phase, and the Technical Program Committee for their high-quality, timely review process. These contributions were crucial to the success of IWSDS 2026.

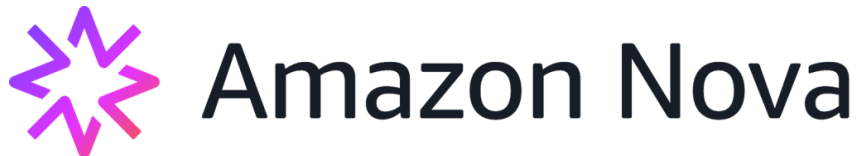
We hope you enjoy your time in Trento!

Giuseppe Riccardi (University of Trento) – IWSDS 2026 General Chair  
Seyed Mahed Mousavi (University of Trento) – IWSDS 2026 Program Chair

## Acknowledgements

The organizers want to thank our partners that highly contributed to the organization and promotion of IWSDS. In concrete, we want to thank: University of Trento, Signals and Interactive Systems Labs (SiSLab), Amazon Nova, ELRA, High Optimization Patterns for Operations Excellence (H.O.P.E.), and Tyto Software Engineering.

### Gold Sponsor



### Silver Sponsors



### Bronze Sponsor



### Institutional Partners



# Organizing Committee

## General Chair

- Giuseppe Riccardi, University of Trento

## Program Chair

- (S.) Mahed Mousavi, University of Trento

## Advisory Chairs

- Maria Inés Torres, University of the Basque Country
- Mkoichiro Yoshino, Tokyo Institute of Technology

## Technical Chairs

- Zoraida Callejas University of Granada
- Shammur Absar Chowdhury, Qatar Computing Research Institute
- Yun-Nung (Vivian) Chen, National Taiwan University
- Frédéric Béchet, Aix Marseille Universit
- Joakim Gustafson, KTH Royal Institute of Technology

## Industrial Track Chairs

- Géraldine Damnati, Orange Research
- Alex Papangelis, Apple Health AI

## Publication Chairs

- Luis Fernando D'Haro, THAU, Universidad Politécnica de Madrid, Spain
- John Mendonça, INESC-ID/IST, Portugal

## Publicity Chair

- Raffaella Bernardi, Free University of Bozen-Bolzano

## Special Session Chairs

- Dilek Hakkani-Tür, University of Illinois Urbana-Champaign
- Giuseppe "Pino" Di Fabrizio, Harvard Business School

- Tatsuya Kawahara, Kyoto University

### **Liaison Chairs**

- Haizhou Li, The Chinese University of Hong Kong (Shenzhen)
- Kallirroi Georgila, University of Southern California

### **Sponsorship Chairs**

- Firoj Alam, Qatar Computing Research Institute
- Michael Johnston, Amazon AGI
- Gokhan Tur, University of Illinois Urbana-Champaign

### **Local Organizing Team**

- Simone Alghisi, PhD student
- Edoardo Cecchinato, Master student
- Olha Khomyn, Research assistant
- Christian Moiola, PhD student
- Massimo Rizzoli, PhD student

### **Program Committee**

- Aakash Bhatnagar
- Akishige Yuguchi
- Alessandra Cervone
- Andrei Coman
- Angus Addlesee
- Arantza del Pozo
- Bjoern Schuller
- Casey Kennington
- Catherine Pelachaud
- Cristina Aceta
- Daniel Hernandez Garcia
- David Traum
- David Griol

- Dilek Hakkani-Tur
- Fabrice Lefèvre
- Felix Burkhardt
- Fernando Fernández-Martínez
- Firoj Alam
- Gabriel Skantze
- Giuseppe Di Fabrizio
- Graham Wilcock
- Hiroaki Sugiyama
- Jacopo Staiano
- Jacqueline Brixey
- Jose Quesada
- Joseph Mariani
- Julia Hirschberg
- Justine Cassell
- Kallirroi Georgila
- Kazunori Komatani
- Kheng Hui Yeo
- Khyati Mahajan
- Koichiro Yoshino
- Koji Inoue
- Kotaro Funakoshi
- Kristiina Jokinen
- Leïla Ben Letaifa
- Luis Fernando D'Haro
- Maike Paetzel-Pruesmann
- Maria Koutsombogera
- Maria Ines Torres
- Marilyn Walker
- Massimo Rizzoli
- Michael McTear
- Mikio Nakano

- Mirco Ravanelli
- Naveen Sai Madiraju
- Nicolas Wagner
- Nigel Ward
- Nina Dethlefs
- Oier Lopez de Lacalle
- Oscar Saz
- Rafael E. Banchs
- Ramesh Manuvinakurike
- Raymond Bond
- Ron Artstein
- Ryuichiro Higashinaka
- Satoshi Nakamura
- Sebastian Möller
- Seiya Kawano
- Simon Keizer
- Simone Alghisi
- Sophie Rosset
- Staffan Larsson
- Stefan Ultes
- Stefan Hillmann
- Susanne Burger
- Svetlana Stoyanchev
- Tatsuya Kawahara
- Timo Baumann
- Vladislav Maraev
- Wolfgang Maier
- Yannick Estève
- Yuki Matsuda
- Yuya Chiba
- Zoraida Callejas

## Table of Contents

<i>MAC: A Multi-Agent Framework for Interactive User Clarification in Multi-turn Conversations</i> Emre Can Acikgoz, Jinoh Oh, Joo Hyuk Jeon, Jie Hao, Heng Ji, Dilek Hakkani-Tur, Gokhan Tur, Xiang Li, Chengyuan Ma and Xing Fan . . . . .	1
<i>FlowSwitch: A State-Aware Framework for Workflow Transitions in Adaptive Dialogue Agents</i> Wen Yu Chang, Luning Qiu, Yi-Hung Liu and Yun-Nung Chen . . . . .	18
<i>Personality Expression in Spoken Dialogue Systems: From Text to Speech</i> Kenta Yamamoto and Kazunori Komatani . . . . .	34
<i>Reproducing Proficiency-Conditioned Dialogue Features with Full-duplex Spoken Dialogue Models</i> Takao Obi, Sadahiro Yoshikawa, Mao Saeki, Masaki Eguchi and Yoichi Matsuyama . . . . .	43
<i>Automatic Evaluation of Open-Domain Real Conversations: Combining Encoder-Based, Dialogue-Based Features and Large Language Models Ratings</i> Cristina Conforto López, Marcos Estecha-Goritagoitia, Mario Rodriguez-Cantelar, Ricardo Cordoba and Luis Fernando D'Haro . . . . .	52
<i>Do audio and visual tokenizers capture backchannels?</i> Benoit Favre and Auriane Boudin . . . . .	64
<i>The Context Trap: Why End-to-End Audio Language Models Fail Multi-turn Dialogues</i> Zhi Rui Tam, Wen Yu Chang and Yun-Nung Chen . . . . .	76
<i>Analysing Next Speaker Prediction in Multi-Party Conversation Using Multimodal Large Language Models</i> Taiga Mori, Koji Inoue, Divesh Lala, Keiko Ochi and Tatsuya Kawahara . . . . .	83
<i>Exploring Emotional Nuances in Spoken Dialogue: Dataset Construction and Prediction of Emotional Dialogue Breakdown</i> Hyuga Nakaguro and Koichiro Yoshino . . . . .	95
<i>Effects of Dialogue Corpora Properties on Fine-Tuning a Moshi-Based Spoken Dialogue Model</i> Yuto Abe, Mao Saeki, Atsumoto Ohashi, Shinnosuke Takamichi, Shiyna Fujie, Tetsunori Kobayashi, Tetsuji Ogawa and Ryuichiro Higashinaka . . . . .	104
<i>Mixed-Initiative Dialogue Management for Human-Virtual Agents Interaction in Forum Theatre Inspired Training</i> Samuel Otofa, Yacine Zerenini, Frederic Bechet, Benoit Favre, Jean-Marie Pergandi and Magalie Ochs . . . . .	109
<i>Analyzing Utterance Selection for Unnoticeable Topic Induction in Target-Guided Conversation Systems</i> Kai Yoshida and Koichiro Yoshino . . . . .	114
<i>Development of an Evaluation System for a Fan-Engagement Chat Application Using LLM-as-a-Judge</i> Yuki Fujita, Yasunobu Sasaki, Ryota Arashi, Hokuto Ootake and Shinya Takahashi . . . . .	123
<i>A Dialogue Agent to Let Users Experience and Gently Enhance the "Gyaru-Mind"</i> Momoka Ikegami, Takuya Kato, Saizo Aoyagi and Tatsunori Hirai . . . . .	128

<i>Towards a proactive cooking companion for the elderly</i> Katarina Esteve, Morgan Fredriksson, Joakim Gustafson, Dimosthenis Kontogiorgos and Timo Mashiyi-Veikkola .....	134
<i>Conversational AI for Virtual Standardized Patients using a Speech-to-Speech LLM</i> Andrew Emerson, Keelan Evanini, Su Somay, Kevin Frome, Le An Ha and Polina Harik .....	142
<i>Can Small-Scale LLMs Balance Content Accuracy and Speaker Faithfulness in Noisy French Dialogue Summarization?</i> Rim Abrougui, Guillaume Lechien, Elisabeth Savatier and Benoît Laurent .....	153
<i>ORCHESTRA: AI-Driven Microservices Architecture to Create Personalized Experiences</i> Jaime Bellver, Samuel Ramos-Varela, Anmol Guragain, Ricardo Córdoba and Luis Fernando D'Haro	158
<i>Benchmarking Multilingual Temporal Reasoning in LLMs: The Temporal Reasoning Dataset</i> Vittorio Mazzia, Sandro Pollastrini, Davide Bernardi, Chiara Rubagotti and Daniele Amberti ..	168
<i>Retrospective Speech Recognition for Spoken Dialogue System: Exploiting Subsequent Utterances to Enhance ASR Performance</i> Ryu Takeda and Kazunori Komatani .....	182
<i>From Fact to Judgment: Investigating the Impact of Task Framing on LLM Conviction in Dialogue Systems</i> Parisa Rabbani, Nimet Beyza Bozdog and Dilek Hakkani-Tur .....	193
<i>Minimal Clips, Maximum Salience: Long Video Summarization via Key Moment Extraction</i> Galann Pennec, Zhengyuan Liu, Nicholas Asher, Philippe Muller and Nancy Chen .....	205
<i>Multilingual and Continuous Backchannel Prediction: A Cross-lingual Study</i> Koji Inoue, Mikey Elmers, Yahui Fu, Zi Haur Pang, Taiga Mori, Divesh Lala, Keiko Ochi and Tatsuya Kawahara .....	222
<i>Vanishing point of attention: A platform for adaptive driver dialogue experiments</i> Morgan Fredriksson, Yanis Yaici, Kevin Lam, Jürgen Königsmann and Jens Edlund .....	231
<i>When social robots see our sketches: evaluating human perception of a robot and a VLM model performance in a drawing task</i> Viktoria Paraskevi Daniilidou, Nikolai Ilinykh and Vladislav Maraev .....	239
<i>Adding Determinism to a Dialogue Agent for a Robotic Environment</i> Oihana Garcia Anakabe, Riccardo Cocola and Cristina Aceta .....	253
<i>Context-Aware Language Understanding in Human-Robot Dialogue with LLMs</i> Svetlana Stoyanchev, Youmna Farag, Simon Keizer, Mohan Li and Rama Sanand Doddipatla ..	262
<i>Learning Vision-Language Alignment in Unified LLMs with 24 Text Tokens per Image</i> Nicola Irmiger, Yixuan Xu, Raphael Kreft, Aram Davtyan, Manuel Kaufmann and Imanol Schlag	275
<i>Incorporating Respect into LLM-Based Academic Feedback: A BI-R Framework for Instructing Students after Q&amp;A Sessions</i> Mayuko Aiba, Daisuke Saito and Nobuaki Minematsu .....	288

<i>The Complementary Role of Para-linguistic cues for Robust Pronunciation Assessment</i> Yassine EL KHEIR, Shammur Absar Chowdhury and Ahmed Ali .....	302
<i>Evaluating LLM Style Transfer Through Readability-Based Age Assessments</i> Maria Di Maro, Antonio Origlia, Leonilda Bilo, Roberta Meo, Pietro Maturi and Francesca Nappo	307
<i>SpeakRL: Synergizing Reasoning, Speaking, and Acting in Language Models with Reinforcement Learning</i> Emre Can Acikgoz, Jinoh Oh, Jie Hao, Joo Hyuk Jeon, Heng Ji, Dilek Hakkani-Tur, Gokhan Tur, Xiang Li, Chengyuan Ma and Xing Fan .....	312
<i>Adaptive Multimodal Sentiment Analysis with Stream-Based Active Learning for Spoken Dialogue Systems</i> Atsuto Ajichi, Takato Hayashi, Kazunori Komatani and Shogo Okada.....	326
<i>Predicting Turn-Taking in Child–Adult Conversations Using Voice Activity Projection</i> Youcef BRAHIMI, César BLANC and Abdellah Fourtassi .....	338
<i>Supporting human operators during customer service interactions with agentic-RAG</i> Juan Barrionuevo-Valenzuela, Daniel Calderón-González, Zoraida Callejas and David Griol ..	348
<i>Analysis of Child-Caregiver Interactions for Developing a Caregiver Spoken Dialogue System</i> Sanae Yamashita, Shota Mochizuki, Yuko Kuma, Ray Sakai, Ayaka Sasaki and Ryuichiro Higashinaka.....	357
<i>Can code-switching improve the user experience with a dialogue system app for recording endangered languages?</i> Jacqueline Brixey and David Traum .....	369
<i>Estimating Relationships between Participants in Multi-Party Chat Corpus</i> Akane Fukushige, Koji Inoue, Keiko Ochi, Tatsuya Kawahara, Sanae Yamashita and Ryuichiro Higashinaka .....	379
<i>WER is Unaware: Assessing How ASR Errors Distort Clinical Understanding in Patient Facing Dialogue</i> Zachary Ellis, Jared Joselowitz, Yash Deo, Yajie Vera He, Anna Kalygina, Aisling Higham, Mana Rahimzadeh, Yan Jia, Ibrahim Habli and Ernest Lim .....	391
<i>ReflectOR: an LLM-based Agent for Post-Operative Surgical Debriefing</i> Lorenzo Fumi, Marco Bombieri, Sara Allievi, Stefano Bonvini, Theodora Chaspari, Marco A. Zenati and Paolo Giorgini .....	418
<i>Detecting Mental Manipulation in Speech via Synthetic Multi-Speaker Dialogue</i> Run Chen, Wen Liang, Ziwei Gong, Lin Ai and Julia Hirschberg .....	428
<i>CoVaPh: A Vision-Language Multi-Agent Dialogue System for Tool-Augmented Pharmacogenetic Reasoning and Personalized Guidance</i> Shang-Chun Luke Lu, Hsin Yang, Hui-Hsin Xue, Ping Lin Tsai, Yu Jing Weng, Shiou-Chi Li, Jen-Wei Huang and Hui Hua Chang .....	441



# Conference Program

**Thursday, February 26, 2026**

**19:00–21:00** Welcome Aperitif

**Friday, February 27, 2026**

**9:15–9:45** [Session 1]: Welcome

**9:45–10:45** [Session 2]: Keynote: Asli Celikyilmaz

**11:15–12:45** [Session 3]: Dialogue and Interactive Systems

11:15–11:30 *MAC: A Multi-Agent Framework for Interactive User Clarification in Multi-turn Conversations*

Emre Can Acikgoz, Jino Oh, Joo Hyuk Jeon, Jie Hao, Heng Ji, Dilek Hakkani-Tur, Gokhan Tur, Xiang Li, Chengyuan Ma and Xing Fan

11:30–11:45 *FlowSwitch: A State-Aware Framework for Workflow Transitions in Adaptive Dialogue Agents*

Wen Yu Chang, Luning Qiu, Yi-Hung Liu and Yun-Nung Chen

11:45–12:00 *Personality Expression in Spoken Dialogue Systems: From Text to Speech*

Kenta Yamamoto and Kazunori Komatani

12:00–12:15 *Reproducing Proficiency-Conditioned Dialogue Features with Full-duplex Spoken Dialogue Models*

Takao Obi, Sadahiro Yoshikawa, Mao Saeki, Masaki Eguchi and Yoichi Matsuyama

12:15–12:30 *Automatic Evaluation of Open-Domain Real Conversations: Combining Encoder-Based, Dialogue-Based Features and Large Language Models Ratings*

Cristina Conforto López, Marcos Estecha-Goritagoitia, Mario Rodriguez-Cantelar, Ricardo Cordoba and Luis Fernando D’Haro

Friday, February 27, 2026 (continued)

**14:00–15:15 [Session 4]: Special Session: Human-Machine Dialogue in the Era of Multimodal Foundation Models**

14:00–14:15 *Do audio and visual tokenizers capture backchannels?*  
Benoit Favre and Auriane Boudin

14:15–14:30 *The Context Trap: Why End-to-End Audio Language Models Fail Multi-turn Dialogues*  
Zhi Rui Tam, Wen Yu Chang and Yun-Nung Chen

14:30–14:45 *Analysing Next Speaker Prediction in Multi-Party Conversation Using Multimodal Large Language Models*  
Taiga Mori, Koji Inoue, Divesh Lala, Keiko Ochi and Tatsuya Kawahara

14:45–15:00 *Exploring Emotional Nuances in Spoken Dialogue: Dataset Construction and Prediction of Emotional Dialogue Breakdown*  
Hyuga Nakaguro and Koichiro Yoshino

**15:15–16:15 [Session 5]: Panel: Human-Machine Dialogue in the Era of Multimodal Foundation Models**

*Panelist*

Yun-Nung (Vivian) Chen - National Taiwan University, Taiwan

Geraldine Damnati - Orange S.A., France

Asli Celikyilmaz - Meta Fundamentals AI Research (FAIR)

Michael Johnston - Amazon

**Friday, February 27, 2026 (continued)**

**16:45–17:30 [Session 6]: Poster Session**

16:45–17:30 *Effects of Dialogue Corpora Properties on Fine-Tuning a Moshi-Based Spoken Dialogue Model*

Yuto Abe, Mao Saeki, Atsumoto Ohashi, Shinnosuke Takamichi, Shiyna Fujie, Tetsunori Kobayashi, Tetsuji Ogawa and Ryuichiro Higashinaka

16:45–17:30 *Mixed-Initiative Dialogue Management for Human-Virtual Agents Interaction in Forum Theatre Inspired Training*

Samuel Ofofa, Yacine Zerenini, Frederic Bechet, Benoit Favre, Magalie Ochs and Jean-Marie Pergandi

16:45–17:30 *Analyzing Utterance Selection for Unnoticeable Topic Induction in Target-Guided Conversation Systems*

Kai Yoshida and Koichiro Yoshino

16:45–17:30 *Development of an Evaluation System for a Fan-Engagement Chat Application Using LLM-as-a-Judge*

Yuki Fujita, Yasunobu Sasaki, Ryota Arashi, Hokuto Ototake and Shinya Takahashi

16:45–17:30 *A Dialogue Agent to Let Users Experience and Gently Enhance the "Gyaru-Mind"*

Momoka Ikegami, Takuya Kato, Saizo Aoyagi and Tatsunori Hirai

16:45–17:30 *Towards a proactive cooking companion for the elderly*

Katarina Esteve, Morgan Fredriksson, joakim gustafson, Dimosthenis Kontogiorgos and Timo Mashiyi-Veikkola

**Saturday, February 28, 2026**

**9:00–10:00 [Session 7]: Keynote: Gabriel Skantze**

**Saturday, February 28, 2026 (continued)**

**10:00–10:45 [Session 8]: Industry Track**

10:00–10:15 *Conversational AI for Virtual Standardized Patients using a Speech-to-Speech LLM*  
Andrew Emerson, Keelan Evanini, Su Somay, Kevin Frome, Le An Ha and Polina Harik

10:15–10:30 *Can Small-Scale LLMs Balance Content Accuracy and Speaker Faithfulness in Noisy French Dialogue Summarization?*  
Rim Abrougui, Guillaume Lechien, Elisabeth Savatier and Benoît Laurent

10:30–10:45 *ORCHESTRA: AI-Driven Microservices Architecture to Create Personalized Experiences*  
Jaime Bellver, Samuel Ramos-Varela, Anmol Guragain, Ricardo Córdoba and Luis Fernando D’Haro

**11:15–12:45 [Session 9]: Resources and Evaluation**

11:15–11:30 *Benchmarking Multilingual Temporal Reasoning in LLMs: The Temporal Reasoning Dataset*  
Vittorio Mazzia, Sandro Pollastrini, Davide Bernardi, Chiara Rubagotti and Daniele Amberti

11:30–11:45 *Retrospective Speech Recognition for Spoken Dialogue System: Exploiting Subsequent Utterances to Enhance ASR Performance*  
Ryu Takeda and Kazunori Komatani

11:45–12:00 *From Fact to Judgment: Investigating the Impact of Task Framing on LLM Conviction in Dialogue Systems*  
Parisa Rabbani, Nimet Beyza Bozdog and Dilek Hakkani-Tur

12:00–12:15 *Minimal Clips, Maximum Saliency: Long Video Summarization via Key Moment Extraction*  
Galann Pennec, Zhengyuan Liu, Nicholas Asher, Philippe Muller and Nancy Chen

12:15–12:30 *Multilingual and Continuous Backchannel Prediction: A Cross-lingual Study*  
Koji Inoue, Mikey Elmers, Yahui Fu, Zi Haur Pang, Taiga Mori, Divesh Lala, Keiko Ochi and Tatsuya Kawahara

**Saturday, February 28, 2026 (continued)**

**14:00–15:15 [Session 10]: Special Session: ConvAI Application in Robotics and Virtual Reality**

14:00–14:15 *Vanishing point of attention: A platform for adaptive driver dialogue experiments*  
Morgan Fredriksson, Yanis Yaici, Kevin Lam, Jurgen Konigsmann and Jens Edlund

14:15–14:30 *When social robots see our sketches: evaluating human perception of a robot and a VLM model performance in a drawing task*  
Viktoria Paraskevi Daniilidou, Nikolai Ilinykh and Vladislav Maraev

14:30–14:45 *Adding Determinism to a Dialogue Agent for a Robotic Environment*  
Oihana Garcia Anakabe, Riccardo Cocola and Cristina Aceta

14:45–15:00 *Context-Aware Language Understanding in Human-Robot Dialogue with LLMs*  
Svetlana Stoyanchev, Youmna Farag, Simon Keizer, Mohan Li and Rama Sanand Doddipatla

**15:15–16:15 [Session 11]: Panel: ConvAI Application in Robotics and Virtual Reality**

*Panelist*

Koichiro Yoshino - Tokyo Institute of Technology / RIKEN, Japan

Gabriel Skantze - KTH Royal Institute of Technology, Sweden

Giulio Jacucci - University of Helsinki, Finland

**Saturday, February 28, 2026 (continued)**

**16:45–17:30 [Session 12]: Poster Session**

16:45–17:30 *Learning Vision-Language Alignment in Unified LLMs with 24 Text Tokens per Image*

Nicola Irmiger, Yixuan Xu, Raphael Kreft, Aram Davtyan, Manuel Kaufmann and Imanol Schlag

16:45–17:30 *Incorporating Respect into LLM-Based Academic Feedback: A BI-R Framework for Instructing Students after Q&A Sessions*

Mayuko Aiba, Daisuke Saito and Nobuaki Minematsu

16:45–17:30 *The Complementary Role of Para-linguistic cues for Robust Pronunciation Assessment*

Yassine EL KHEIR, Shammur Absar Chowdhury and Ahmed Ali

16:45–17:30 *Evaluating LLM Style Transfer Through Readability-Based Age Assessments*

Maria Di Maro, Antonio Origlia, Leonilda Bilo, Roberta Meo, Pietro Maturi and Francesca Nappo

16:45–17:30 *SpeakRL: Synergizing Reasoning, Speaking, and Acting in Language Models with Reinforcement Learning*

Emre Can Acikgoz, Jinoh Oh, Jie Hao, Joo Hyuk Jeon, Heng Ji, Dilek Hakkani-Tur, Gokhan Tur, Xiang Li, Chengyuan Ma and Xing Fan

16:45–17:30 *Adaptive Multimodal Sentiment Analysis with Stream-Based Active Learning for Spoken Dialogue Systems*

Atsuto Ajichi, Takato Hayashi, Kazunori Komatani and Shogo Okada

**Sunday, March 1, 2026**

**9:00–10:00 [Session 13]: Keynote: Giulio Jacucci**

**Sunday, March 1, 2026 (continued)**

**10:00–10:45 [Session 14]: Industry Talk**

*AI in a real use case of TP Italy Group: results achieved and new challenges*

Vincenzo Giliberti & Vincenzo Lanzolla

*New Avenues in Dialog: Coding and Adversarial Security*

Michael Johnston, Amazon

**11:15–12:45 [Session 15]: Human-centered Interaction**

11:15–11:30 *Predicting Turn-Taking in Child–Adult Conversations Using Voice Activity Projection*

Youcef BRAHIMI, César BLANC and Abdellah Fourtassi

11:30–11:45 *Supporting human operators during customer service interactions with agentic-RAG*

Juan Barrionuevo-Valenzuela, Daniel Calderón-González, Zoraida Callejas and David Griol

11:45–12:00 *Analysis of Child-Caregiver Interactions for Developing a Caregiver Spoken Dialogue System*

Sanae Yamashita, Shota Mochizuki, Yuko Kuma, Ray Sakai, Ayaka Sasaki and Ryuichiro Higashinaka

12:00–12:15 *Can code-switching improve the user experience with a dialogue system app for recording endangered languages?*

Jacqueline Brixey and David Traum

12:30–12:45 *Estimating Relationships between Participants in Multi-Party Chat Corpus*

Akane Fukushige, Koji Inoue, Keiko Ochi, Tatsuya Kawahara, Sanae Yamashita and Ryuichiro Higashinaka

**Sunday, March 1, 2026 (continued)**

**14:00–15:15 [Session 16] Special Session: ConvAI in the Health Domain**

14:00–14:15 *WER is Unaware: Assessing How ASR Errors Distort Clinical Understanding in Patient Facing Dialogue*

Zachary Ellis, Jared Joselowitz, Yash Deo, Yajie Vera He, Anna Kalygina, Aisling Higham, Mana Rahimzadeh, Yan Jia, Ibrahim Habli and Ernest Lim

14:15–14:30 *ReflectOR: an LLM-based Agent for Post-Operative Surgical Debriefing*

Lorenzo Fumi, Marco Bombieri, Sara Allievi, Stefano Bonvini, Theodora Chaspari, Marco A. Zenati and Paolo Giorgini

14:45–15:00 *Detecting Mental Manipulation in Speech via Synthetic Multi-Speaker Dialogue*

Run Chen, Wen Liang, Ziwei Gong, Lin Ai and Julia Hirschberg

15:00–15:15 *CoVaPh: A Vision-Language Multi-Agent Dialogue System for Tool-Augmented Pharmacogenetic Reasoning and Personalized Guidance*

Shang-Chun Luke Lu, Hsin Yang, Hui-Hsin Xue, Ping Lin Tsai, Yu Jing Weng, Shiou-Chi Li, Jen-Wei Huang and Hui Hua Chang

**15:15–16:15 [Session 17] Panel: ConvAI in the Health Domain**

*Panelist*

Tatsuya Kawahara - Kyoto University, Japan

Marco Zenati - Harvard Medical School, USA

Giuseppe Riccardi - University of Trento, Italy

Tommaso Ciulli - Secretary of the Italian Society of Digital Psychology

**16:45–17:30 [Session 18] Paper Awards and Closing Remarks**