

SIGDIAL 2026



**The 27th Annual Meeting of the  
Special Interest Group on Discourse and Dialogue**

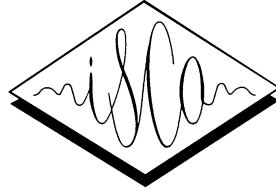
**Proceedings of the Conference**

August 2 - 5, 2026  
Emory University, Atlanta, Georgia, USA

In cooperation with:  
Association for Computational Linguistics (ACL)



International Speech Communication Association (ISCA)



We thank our sponsor:  
Department of Computer Science  
Emory University

©2026 SIGDIAL

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 979-8-89176-438-5

## Message from the General Chair and Program Chairs

We are delighted to present the proceedings of SIGDIAL 2026, the 27th Annual Meeting of the Special Interest Group on Discourse and Dialogue. SIGDIAL has established itself as a premier venue for the dissemination and discussion of cutting-edge research in discourse and dialogue.

This year, the conference was held as a fully in-person event from August 2 to 5, 2026, at Emory University in Atlanta, Georgia, USA. The SIGDIAL 2026 program featured a rich and diverse set of activities: three keynote talks, five oral presentation sessions, three poster sessions, a business meeting, an invited talk, and a social event.

In addition, two satellite workshops took place on the day preceding the main conference: the Workshop on Chatbots and Agentic Technologies (WOCHAT 2026) and the Workshop on Spoken Dialogue Systems for PhDs, PostDocs & New Researchers (YRRSDS 2026).

SIGDIAL 2026 received 160 submissions, of which 134 remained after desk rejections and withdrawals, including 129 submitted via Softconf/START and 5 received via the ACL Rolling Review (ARR). The review process was supported by 15 Senior Program Committee (SPC) members, who led discussions and provided meta-reviews, alongside 133 Program Committee (PC) reviewers. Each submission received at least three detailed reviews. Final decisions were made through careful consideration of the reviews, meta-reviews, and discussion outcomes, facilitated by the SPCs.

We are deeply grateful to all members of the Program Committee and Senior Program Committee for their thoughtful evaluations and dedication. Their contributions were vital in assembling a high-quality technical program. Aiming for both quality and diversity, a total of 61 papers were accepted: 55 long papers, 4 short papers, and 2 demonstration papers, yielding an overall acceptance rate of 45.5%. The selected contributions reflect the broad and vibrant landscape of current research in discourse and dialogue.

During this edition, we were honoured to welcome three distinguished keynote speakers:

- Prof. Larry Heck, Professor of Electrical and Computer Engineering and Interactive Computing at the Georgia Institute of Technology, with a talk entitled “The Conversation Inside the Model: Probing and Controlling Interactional Structure in Multimodal Foundation Models”
- Prof. Nanyun (Violet) Peng, Associate Professor of Computer Science at the University of California, Los Angeles, presenting “David and Goliath: Compute-Efficient Strategies for LLM Steering”
- Prof. Dilek Hakkani-Tür, Professor of Computer Science at the University of Illinois at Urbana-Champaign, delivering “Towards Safe and Self-Evolving Conversational Agents”

We were also pleased to host an invited talk by Dr. Matthew Marge, Program Manager at the Defense Advanced Research Projects Agency (DARPA), entitled “Frontier AI for National Security: The AI Forge Program”.

We would like to thank the SIGDIAL officers and Emory University for their continued support and guidance. We are also immensely appreciative of the tireless work of the organizing committee:

- Local Chairs: Kartik Goyal (Georgia Institute of Technology, USA) and Sandeep Soni (Emory University, USA)

- Sponsorship Chair: Teruhisa Misu (Honda Research Institute, USA)
- Publication Chair: Ali Emami (Emory University, USA)

We gratefully acknowledge the generous support of our sponsor: the Department of Computer Science at Emory University.

Finally, it was our great pleasure to welcome all the participants to Atlanta and Emory University. A dynamic Southern capital where history meets innovation, Atlanta is a vibrant hub of civil rights heritage, world-class cuisine, thriving arts, and business, all beneath a skyline of soaring towers and lush urban canopy. Emory University, a leading research university renowned for its contributions to healthcare, liberal arts, and global scholarship, hosted SIGDIAL 2026 at its Psychology and Interdisciplinary Sciences (PAIS) Building, providing participants with not only a scientifically enriching experience but also a warm introduction to the culture and hospitality of the American South.

We hope everyone enjoyed a productive and memorable conference experience and left with lasting impressions of SIGDIAL 2026 as a successful event.

Sincerely,

General Chair: Jinho D. Choi (Emory University, USA)

Program Co-Chairs: Yun-Nung (Vivian) Chen (National Taiwan University, Taiwan) and Kotaro Funakoshi (Institute of Science Tokyo, Japan)



# **Organizing Committee**

## **General Chair**

Jinho D. Choi, Emory University, USA

## **Program Chairs**

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

Kotaro Funakoshi, Institute of Science Tokyo, Japan

## **Local Chairs**

Kartik Goyal, Georgia Institute of Technology, USA

Sandeep Soni, Emory University, USA

## **Sponsorship Chair**

Teruhisa Misu, Honda Research Institute, USA

## **Publication Chair**

Ali Emami, Emory University, USA

## Program Committee

### Senior Program Committee:

Nicholas Asher (CNRS Institut de Recherche en Informatique de Toulouse)  
Frederic Bechet (Aix Marseille Universite - LIS/CNRS)  
Luis Fernando D'Haro (Speech Technology and Machine Learning Group, E.T.S.I. Telecomunicación, IPTC, Universidad Politécnica de Madrid)  
Kristiina Jokinen (AIRC, AIST and University of Helsinki)  
Staffan Larsson (University of Gothenburg)  
Fabrice Lefèvre (Avignon Univ.)  
Pierre Lison (Norwegian Computing Centre)  
Mikio Nakano (C4A Research Institute, Inc./Univ. of Osaka/Nagoya Univ./Nagoya Inst. of Tech.)  
Gabriel Skantze (KTH Speech Music and Hearing)  
Svetlana Stoyanchev (Toshiba Europe)  
David Traum (University of Southern California Institute for Creative Technologies)  
Gokhan Tur (University of Illinois, Urbana Champaign)  
Marilyn Walker (University of California Santa Cruz)  
Koichiro Yoshino (The University of Osaka / Institute of Science Tokyo / RIKEN GRP / Nara Institute of Science and Technology)  
Frances Yung (Saarland University)

### Program Committee:

Alafate Abulimiti  
Emre Can Acikgoz  
Jan Alexandersson  
Ron Artstein  
Masayuki Asahara  
Timo Baumann  
Nitu Bharati  
Tom Bourgeade  
Kevin Bowden  
Hendrik Buschmeier  
Jie Cao  
Giuseppe Carenini  
Justine Cassell  
Alberto Cetoli  
Hsien-Tsung Chang  
Akshay Chaturvedi  
Hui Chen  
Yuya Chiba  
Gérard Chollet  
Tahiya Chowdhury  
Jaroslaw Chudziak  
Gaël de Chalendar  
Nina Dethlefs  
David DeVault  
Bosheng Ding  
Rama Sanand Doddipatla



Cecilia Domingo  
Vardhan Dongre  
Ondrej Dusek  
Ali Emami  
Keelan Evanini  
Effat Farhana  
Mary Ellen Foster  
Yahui Fu  
Yingxue Fu  
Shinya Fujie  
Kallirroi Georgila  
Felix Gervits  
Emer Gilmartin  
Jonathan Ginzburg  
Prakhar Gupta  
Joakim Gustafson  
Dilek Hakkani-Tur  
Jie He  
Michael Heck  
Behnam Hedayatnia  
Ryuichiro Higashinaka  
Jessica Huynh  
Takamasa Iio  
Michimasa Inaba  
Mert Inan  
Koji Inoue  
Naoto Iwahashi  
Shinji Iwata  
Tatsuya Kawahara  
Seiya Kawano  
Casey Kennington  
Takyoung Kim  
Xenia Klinge  
Kazunori Komatani  
Mario Kubek  
Gaurav Kumar  
Divesh Lala  
Ekaterina Lapshinova-Koltunski  
Kornel Laskowski  
Hae-Na Lee  
Chuyuan Li  
Dongqi Liu  
Wolfgang Maier  
Vladislav Maraev  
Collin McMillan  
Yasuhiro Minami  
Teruhisa Misu  
Yusuke Miyao  
Kouki Miyazawa  
Hiroki Mori  
Fatema Nafa

Kouta Nakayama  
Hiromi Narimatsu  
Anna Nedoluzhko  
Noriki Nishida  
Douglas O'Shaughnessy  
Shogo Okada  
Michael Orme  
Suraj Pandey  
Patrick Paroubek  
Sachin Pawar  
Massimo Poesio  
Jeanne Powell  
Matthew Purver  
Zhiyang Qi  
Kun Qian  
Parisa Rabbani  
Lina M. Rojas Barahona  
Antonio Roque  
Rimvydas Rubavicius  
Ramon Ruiz-Dolz  
Rikunari Sagara  
Saurav Sahay  
Yuki Saito  
Shiki Sato  
Yoshinao Sato  
Phillip Schneider  
Samira Shaikh  
Anthony Sicilia  
Carl Strathearn  
Hiroaki Sugiyama  
Yi-Jyun Sun  
Ekaterina Svikhnushina  
Ryu Takeda  
António Teixeira  
Mariat Theune  
Kate Thompson  
Ryoko Tokuhisa  
Maria Ines Torres  
Suraj Tripathi  
Kazuya Tsubokura  
Yuiko Tsunomori  
Markku Turunen  
David Vandyke  
Nicolas Wagner  
Nicholas Walker  
Yi-Chia Wang  
Nigel Ward  
Graham Wilcock  
Win Pa Pa  
Moritz Wolf  
Kenta Yamamoto

Xifeng Yan  
W. L. Yeung  
Yanchao Yu  
Amir Zeldes  
Ingrid Zukerman



## Table of Contents

|                                                                                                                                                                                                                                                                             |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>DialogueSidon: Recovering Full-Duplex Dialogue Tracks from In-the-Wild Dialogue Audio</i><br>Wataru Nakata, Yuki Saito, Kazuki Yamauchi, Emiru Tsunoo and Hiroshi Saruwatari .....                                                                                       | 1   |
| <i>Rethinking Binary Evaluation of Turn-Taking under Inherent Ambiguity</i><br>Yunosuke Kubo, Kenta Yamamoto, Ryu Takeda and Kazunori Komatani .....                                                                                                                        | 13  |
| <i>Tinker Tales: A Tangible Dialogue System for Child–AI Co-Creative Storytelling</i><br>Nayoung Choi, Jiseung Hong, Peace Cyebukayire, Ikseon Choi and Jinho D. Choi .....                                                                                                 | 25  |
| <i>Suppressing Unnecessary Clarification Requests for Unknown Word Acquisition in Spoken Dialogue Using Syllable-Based ASR Confidence</i><br>Takumi Furuta, Ryu Takeda and Kazunori Komatani .....                                                                          | 51  |
| <i>Towards Conversational Patient History-Taking: Voice-Interactive AI Agents for Pre-visit Dementia Diagnostic Interviews</i><br>Andrew G. Breithaupt, Nayoung Choi, James D. Finch, Jeanne M. Powell, Arin L. Nelson, Oz A. Alon, Howard J. Rosen and Jinho D. Choi ..... | 62  |
| <i>ClarVis: A Dataset of Clarification Requests and Grounding in Collaborative Data Visualization Dialogues</i><br>Abari Bhattacharya and Barbara Di Eugenio .....                                                                                                          | 79  |
| <i>Accuracy and Satisfaction in Multi-Turn LLM Dialogues for NFR Assessment</i><br>ali pourghasemi fatideh, Wilder Baldwin, Maria Dhakal, Collin McMillan and Sepideh Ghanavati                                                                                             | 91  |
| <i>Breaking the Script: Do Role-Playing Agents Maintain Goal Alignment under Distraction?</i><br>Dongxu Lu, Albert Gatt and Johan Jeuring .....                                                                                                                             | 107 |
| <i>ArgAssist: LLM-based Argument Synthesis for Insurance Disputes</i><br>Anubhav Sinha, Nitin Ramrakhiyani, Sachin Pawar, Isha Narang and Manoj Apte .....                                                                                                                  | 124 |
| <i>Detecting Experiential Intertextuality Across Migration Routes: Beyond Surface Similarity in French Narratives</i><br>Sakayo Toadoun Sari, Nelly Robin, Michelle Auzanneau, Lakhdar Sais, Véronique Petit, Marie Veniard, Said Jabbour and Fabien Delorme .....          | 139 |
| <i>Conversational Grounding in Large Language Models: Evaluation Methods, Challenges and Future Directions</i><br>Michelle Elizabeth, Gwénolé Lecorvé, Lina M. Rojas Barahona and Magalie Ochs .....                                                                        | 151 |
| <i>Sensor-Augmented Voice Activity Projection for Enhancing Turn-Taking Prediction</i><br>Satoki Hamanaka, Yasue Kishino, Yuiko Tsunomori, Shin Mizutani, Yuya Chiba, Tadashi Okoshi and Jin Nakazawa .....                                                                 | 164 |
| <i>A Dialogue System for Second Language Learning with Dynamic Adaptation to In-Dialogue Difficulty Changes</i><br>Taku Morioka, Junya Takayama and Tomoyuki Kajiwara .....                                                                                                 | 170 |
| <i>Ping-Ponder: Concurrent Dual-Agent Spoken Dialogue via Shared Belief State</i><br>Akiko Masaki-Kato and Gulzar Haris .....                                                                                                                                               | 188 |

|                                                                                                                                   |     |
|-----------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Rethinking Role-Playing Evaluation: Anonymous Benchmarking and A Systematic Study of Personality Effects</i>                   |     |
| Ji-Lun Peng and Yun-Nung Chen . . . . .                                                                                           | 205 |
| <i>I Understand How You Feel: Enhancing Deeper Emotional Support Through Multilingual Emotional Validation in Dialogue System</i> |     |
| Zi Haur Pang, Yahui Fu, Koji Inoue and Tatsuya Kawahara . . . . .                                                                 | 219 |
| <i>For What Reason? Interpreting Models' Encoding of Causation and Antithesis</i>                                                 |     |
| Abhidip Bhattacharyya and Shira Wein . . . . .                                                                                    | 237 |
| <i>CrossOracle: An Agentic Framework for Real-Time Expert Personas with Parallelized Acoustic Synthesis</i>                       |     |
| yash raj singh . . . . .                                                                                                          | 254 |
| <i>From Felt Sense to Words: Construction and Analysis of a Focusing Dialogue Dataset for Verbalization</i>                       |     |
| Yuiko Tsunomori, Yuya Chiba and Yosuke Koshikawa . . . . .                                                                        | 258 |
| <i>Retrieving Responses Useful as References in Counseling via LLM-Generated Structurally Analogous Dialogues</i>                 |     |
| Yu Nakagawa, Nozomu Ikeda, Kotaro Funakoshi and Manabu Okumura . . . . .                                                          | 271 |
| <i>Embodied Multi-Agent Coordination by Aligning World Models Through Dialogue</i>                                                |     |
| Vardhan Dongre and Dilek Hakkani-Tur . . . . .                                                                                    | 293 |
| <i>SagaQA: A Multi-hop Reasoning Benchmark for Long-form Narrative Understanding in TV Series</i>                                 |     |
| Galann Pennec, Zhengyuan Liu, Nicholas Asher, Philippe Muller and Nancy Chen . . . . .                                            | 307 |
| <i>How Ethos and Pathos Appeals Resonate in Reader Interpretations of Social Media Messages</i>                                   |     |
| Ewelina Gajewska, Katarzyna Budzynska, Jarosław A. Chudziak and Liesbeth Allein . . . . .                                         | 327 |
| <i>GBC: Gradient-Based Connections for Optimizing Multi-Agent Systems</i>                                                         |     |
| Xiaocheng Yang, Abdulrahman Alrabah, Dilek Hakkani-Tür and Gokhan Tur . . . . .                                                   | 342 |
| <i>A French OSCE Dialogue Dataset and Controllable Virtual Patient System for Clinical Training</i>                               |     |
| Doria Bonzi, Tom Bourgeade, Fabrice Lefèvre and Irina Illina . . . . .                                                            | 357 |
| <i>Do LLMs Use Relational Structure in KG-Grounded Dialogue? Evidence from Entity and Relation Perturbation</i>                   |     |
| Dominic Okonkwo and Ismailcem Budak Arpinar . . . . .                                                                             | 375 |
| <i>DiPS: Dialogue Policy Selection for High-Stakes Persuasion Agents</i>                                                          |     |
| Tianyi Zhang, Mousumi Das, Abrar Anwar, Jesse Thomason and David Traum . . . . .                                                  | 385 |
| <i>From Propositional to Perceptual Asymmetry: Extending FPO to Asymmetric Partial Information Dialogue</i>                       |     |
| Yifan Zhu, Kyeongmin Rim and James Pustejovsky . . . . .                                                                          | 403 |
| <i>Disentangling Annotator Skill from Verifier Strictness in Cross-Verified Dialogue Annotation</i>                               |     |
| Zihao Tao, John A. Prado, Ignazio Steven LaManna, Ryan Puterbaugh, Mim Datta and Julia Hirschberg . . . . .                       | 414 |
| <i>Beyond Supervised Clarification: Input Rewriting with LLMs for Dialogue Discourse Parsing</i>                                  |     |
| Yiming Liu, Ziyue Zhang, Zhichao Xu, Xin Yu, Yingheng Tang, Tianyu Jiang and Jie Cao . . . . .                                    | 423 |
| <i>Weakly Supervised Temporal Modeling of Latent Dynamics in Dyadic Conversations</i>                                             |     |
| Tahiya Chowdhury . . . . .                                                                                                        | 440 |

|                                                                                                                                                  |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>LLMs Out-of-the-Box Do Not Generate Context-Appropriate Word Order in Russian</i>                                                             |     |
| Alina Shabaeva, John Frederick Bailyn and Owen Rambow . . . . .                                                                                  | 453 |
| <i>Chronological Thinking in Full-Duplex Spoken Dialogue Language Models</i>                                                                     |     |
| Donghang Wu, Haoyang Zhang, Chen Chen, Tianyu Zhang, Fei Tian, Xuerui Yang, Gang Yu, Hexin Liu, Nana Hou, Yuchen Hu and Eng Siong Chng . . . . . | 473 |
| <i>Exploring Retrieval Augmented Generation Approaches for Natural Language Question Understanding</i>                                           |     |
| Christoph Kowalski, Amelie Sophie Robrecht-Hilbig, Vincent Emmerling and Stefan Kopp . . .                                                       | 486 |
| <i>Tree-Based Interview Topic Guidance for Collecting Target Information under Adaptive Topic continuation/switching</i>                         |     |
| Fuminori Nagasawa, Ekai Hashimoto and Shun Shiramatsu . . . . .                                                                                  | 497 |
| <i>COCORELI: Enforcing Execution Preconditions for Reliable Collaborative Instruction Following</i>                                              |     |
| Swarnadeep Bhar, Omar Naim, Eleni Metheniti, Loïc Cabannes, Bastien Navarri, Morteza Kamal-addini Ezzabady and Nicholas Asher . . . . .          | 516 |
| <i>StagePilot: Stage-Level Planning for Long-Horizon Dialogue Simulation in Cybergrooming</i>                                                    |     |
| Heajun An, Qi Zhang, Minqian Liu, Xinyi Zhang, Sang Won Lee, Lifu Huang, Pamela Wisniewski and Jin-Hee Cho . . . . .                             | 536 |
| <i>Generation of Dialogue Breakdown Repair Utterance in Non-task Oriented Conversation</i>                                                       |     |
| Kazuya Tsubokura, Yurie Iribe and Norihide Kitaoka . . . . .                                                                                     | 553 |
| <i>From Sound to Symptom: Real-Time Respiratory Signal Understanding for Conversational Healthcare Agents</i>                                    |     |
| Tanmay Laud, Herprit Mahal and Subhabrata Mukherjee . . . . .                                                                                    | 559 |
| <i>Interpretable Satisfaction Modeling in Sales Dialogues: A Metrics-based Approach Validated Against Text-based Model</i>                       |     |
| Natalia Krawczyk, Alicja Kasicka, Bartosz Przybył and Justyna Gromada . . . . .                                                                  | 565 |
| <i>Adaptive Emotion Management in Human-robot Dialogue using Online Group Relative Policy Optimization</i>                                       |     |
| Anna Manaseryan and Casey Kennington . . . . .                                                                                                   | 585 |
| <i>PersonaKit (PK): A Plug-and-Play Platform for User Testing Diverse Roles in Simulated Full-Duplex Dialogue</i>                                |     |
| Hyunbae Jeon and Jinho D. Choi . . . . .                                                                                                         | 597 |
| <i>Question Types for Knowledge Acquisition in LLM-based Dialogue Systems: Experiments with Simulated and Human Users</i>                        |     |
| Kazunori Komatani, Ryu Takeda and Mikio Nakano . . . . .                                                                                         | 602 |
| <i>TRACER: Early Failure Detection for Task-Oriented Dialogue</i>                                                                                |     |
| Erfan Nourbakhsh, Rocky Slavin, Ke Yang and Anthony Rios . . . . .                                                                               | 614 |
| <i>Relex: A Common Denominator for Connectives and Discourse Relations in French</i>                                                             |     |
| Fatou Sow, Yannick Toussaint and Mathieu Constant . . . . .                                                                                      | 636 |
| <i>MemeBuddy: Dialog-Style Audio Representations for Engaging Non-Visual Meme Experiences</i>                                                    |     |
| Chirag Bhansali, Vikas Ashok and Hae-Na Lee . . . . .                                                                                            | 649 |

|                                                                                                                                              |     |
|----------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>On the Structure of Address in Multi-Party Dialogue: From Discrete Labels to Continuous Levels</i>                                        |     |
| Taiga Mori, Koji Inoue, Divesh Lala and Tatsuya Kawahara .....                                                                               | 668 |
| <i>Bridging Talk and Thought: Understanding Dialogue Dynamics Across Collaborative Problem-Solving Contexts</i>                              |     |
| Zhengyuan Liu, Stella Yin, Min-Yen Kan and Nancy Chen .....                                                                                  | 680 |
| <i>Seeing Is Not Sharing: Some Vision-Language Models Overestimate Common Ground in Asymmetric Dialogue</i>                                  |     |
| Nan Li, Albert Gatt and Massimo Poesio .....                                                                                                 | 694 |
| <i>Evaluation of Paralinguistic-Aware Spoken Dialogue Systems using Next-Utterance Classification</i>                                        |     |
| Kouki Miyazawa and Yoshinao Sato .....                                                                                                       | 711 |
| <i>Better Scores, Worse Grounding: Hidden Regressions after Fine-Tuning in Dialogue Fact Verification</i>                                    |     |
| Hyunkyung Park and Arkaitz Zubiaga .....                                                                                                     | 720 |
| <i>LLMs and their Limited Theory of Mind: Evaluating Mental State Annotations in Situated Dialogue</i>                                       |     |
| Katharine Kowalyszyn and Matthias Scheutz .....                                                                                              | 738 |
| <i>Enhancing Operational Safety via Agentic Dialogue Hazard Identification Analysis</i>                                                      |     |
| Sanjay Das, Ran Elgedawy, Ethan Seefried, Ryan Burchfield and Tirthankar Ghosal .....                                                        | 757 |
| <i>Building Task-Oriented Dialogue Systems via Instruction Guidance without Annotated Data</i>                                               |     |
| Henry Gao and Jinho D. Choi .....                                                                                                            | 773 |
| <i>Bounded Conversational Memory with Hybrid Retrieval and Evidence Highlighting for Multi-Session Dialogue Systems</i>                      |     |
| Manasa Mudunuri, Chandra Kiran Guntupalli, Mario Michael Kubek, Shiraj Pokharel and Murray Patterson .....                                   | 786 |
| <i>Too Polite to Disagree: Understanding Sycophancy Propagation in Multi-Agent Systems</i>                                                   |     |
| Vira Kasprova, Amruta Parulekar, Abdulrahman AlRabah, Krishna Agaram, Ritwik Garg, Sagar Jha, Nimet Beyza Bozdog and Dilek Hakkani-Tur ..... | 795 |
| <i>When the Database Fails: Prompting LLM Dialogue Agents for Safe Recovery in Task-Oriented Dialogue</i>                                    |     |
| Mohammad Alijanpour Shalmani, Alale Rezvani Boroujeni and Jiann S. Yuan .....                                                                | 815 |
| <i>MTDiag: A Multi-Turn Diagnostic Dataset Towards Clinically Meaningful LLM Evaluation</i>                                                  |     |
| Pia Chouayfati, Alexander M. Fichtl, Miriam Anschütz, George Doumat and Georg Groh .....                                                     | 821 |
| <i>PERSONAJUDGE: Simulating Individual Human Preference Judgments with Evaluator-Specific Demonstration Data</i>                             |     |
| Zeyu He, Xuan Qi, Subramanian Chidambaram, Zhichao Xu, Vinayak Arannil, Lydia Chilton and Alex C. Williams .....                             | 841 |
| <i>Fractional Decay KV-Cache: Ownership-Aware Memory Management for Improved Inference Relevancy in Dialog Systems</i>                       |     |
| Sukanta Ganguly .....                                                                                                                        | 863 |
| <i>A Reinforcement Learning-Based Facilitator for Simulated Group Motivational Interviewing</i>                                              |     |
| Alafate Abulimiti, Vladislav Maraev, Agnès Helme-Guizon and Catherine Pelachaud .....                                                        | 871 |



# Conference Program

**2026-08-03**

**08:00–09:00**    *Registration*

**09:00–09:30**    *Opening*

**09:30–10:30**    *Keynote 1: Larry Heck*

**10:30–11:00**    *Coffee Break*

**11:00–12:30**    **Session O1: Oral Session 1**

Room: PAIS 290

*DialogueSidon: Recovering Full-Duplex Dialogue Tracks from In-the-Wild Dialogue Audio*

Wataru Nakata, Yuki Saito, Kazuki Yamauchi, Emiru Tsunoo and Hiroshi Saruwatari

*Rethinking Binary Evaluation of Turn-Taking under Inherent Ambiguity*

Yunosuke Kubo, Kenta Yamamoto, Ryu Takeda and Kazunori Komatani

*Tinker Tales: A Tangible Dialogue System for Child–AI Co-Creative Storytelling*

Nayoung Choi, Jiseung Hong, Peace Cyebukayire, Ikseon Choi and Jinho D. Choi

*Suppressing Unnecessary Clarification Requests for Unknown Word Acquisition in Spoken Dialogue Using Syllable-Based ASR Confidence*

Takumi Furuta, Ryu Takeda and Kazunori Komatani

*Towards Conversational Patient History-Taking: Voice-Interactive AI Agents for Pre-visit Dementia Diagnostic Interviews*

Andrew G. Breithaupt, Nayoung Choi, James D. Finch, Jeanne M. Powell, Arin L. Nelson, Oz A. Alon, Howard J. Rosen and Jinho D. Choi

**12:30–14:00**    *Lunch Break*

**2026-08-03 (continued)**

**14:00–15:30    Session O2: Oral Session 2**

Room: PAIS 290

*ClarVis: A Dataset of Clarification Requests and Grounding in Collaborative Data Visualization Dialogues*

Abari Bhattacharya and Barbara Di Eugenio

*Accuracy and Satisfaction in Multi-Turn LLM Dialogues for NFR Assessment*

ali pourghasemi fatideh, Wilder Baldwin, Maria Dhakal, Collin McMillan and Sepideh Ghanavati

*Breaking the Script: Do Role-Playing Agents Maintain Goal Alignment under Distraction?*

Dongxu Lu, Albert Gatt and Johan Jeuring

*ArgAssist: LLM-based Argument Synthesis for Insurance Disputes*

Anubhav Sinha, Nitin Ramrakhiyani, Sachin Pawar, Isha Narang and Manoj Apte

*Detecting Experiential Intertextuality Across Migration Routes: Beyond Surface Similarity in French Narratives*

Sakayo Toadoun Sari, Nelly Robin, Michelle Auzanneau, Lakhdar Sais, Véronique Petit, Marie Veniard, Said Jabbour and Fabien Delorme

**15:30–16:00    Coffee Break**

**16:00–17:30    Session P1: Poster Session 1**

Room: PAIS main lobby

*Conversational Grounding in Large Language Models: Evaluation Methods, Challenges and Future Directions*

Michelle Elizabeth, Gwénolé Lecorvé, Lina M. Rojas Barahona and Magalie Ochs

*Sensor-Augmented Voice Activity Projection for Enhancing Turn-Taking Prediction*

Satoki Hamanaka, Yasue Kishino, Yuiko Tsunomori, Shin Mizutani, Yuya Chiba, Tadashi Okoshi and Jin Nakazawa

*A Dialogue System for Second Language Learning with Dynamic Adaptation to In-Dialogue Difficulty Changes*

Taku Morioka, Junya Takayama and Tomoyuki Kajiwara

*Ping-Ponder: Concurrent Dual-Agent Spoken Dialogue via Shared Belief State*

Akiko Masaki-Kato and Gulzar Haris

2026-08-03 (continued)

*Rethinking Role-Playing Evaluation: Anonymous Benchmarking and A Systematic Study of Personality Effects*

Ji-Lun Peng and Yun-Nung Chen

*I Understand How You Feel: Enhancing Deeper Emotional Support Through Multilingual Emotional Validation in Dialogue System*

Zi Haur Pang, Yahui Fu, Koji Inoue and Tatsuya Kawahara

*For What Reason? Interpreting Models' Encoding of Causation and Antithesis*

Abhidip Bhattacharyya and Shira Wein

*CrossOracle: An Agentic Framework for Real-Time Expert Personas with Parallelized Acoustic Synthesis*

yash raj singh

*From Felt Sense to Words: Construction and Analysis of a Focusing Dialogue Dataset for Verbalization*

Yuiko Tsunomori, Yuya Chiba and Yosuke Koshikawa

*Retrieving Responses Useful as References in Counseling via LLM-Generated Structurally Analogous Dialogues*

Yu Nakagawa, Nozomu Ikeda, Kotaro Funakoshi and Manabu Okumura

*Embodied Multi-Agent Coordination by Aligning World Models Through Dialogue*

Vardhan Dongre and Dilek Hakkani-Tur

*SagaQA: A Multi-hop Reasoning Benchmark for Long-form Narrative Understanding in TV Series*

Galann Pennec, Zhengyuan Liu, Nicholas Asher, Philippe Muller and Nancy Chen

**2026-08-04**

**08:30–09:00**    *Registration*

**09:00–10:00**    *Keynote 2: Nanyun (Violet) Peng*

**10:00–10:30**    *Coffee Break*

**10:30–12:00**    **Session O3: Oral Session 3**  
Room: PAIS 290

*How Ethos and Pathos Appeals Resonate in Reader Interpretations of Social Media Messages*

Ewelina Gajewska, Katarzyna Budzynska, Jarosław A. Chudziak and Liesbeth Allein

*GBC: Gradient-Based Connections for Optimizing Multi-Agent Systems*

Xiaocheng Yang, Abdulrahman Alrabah, Dilek Hakkani-Tür and Gokhan Tur

*A French OSCE Dialogue Dataset and Controllable Virtual Patient System for Clinical Training*

Doria Bonzi, Tom Bourgeade, Fabrice Lefèvre and Irina Illina

*Do LLMs Use Relational Structure in KG-Grounded Dialogue? Evidence from Entity and Relation Perturbation*

Dominic Okonkwo and Ismailcem Budak Arpinar

*DiPS: Dialogue Policy Selection for High-Stakes Persuasion Agents*

Tianyi Zhang, Mousumi Das, Abrar Anwar, Jesse Thomason and David Traum

**12:00–13:30**    *Lunch Break*

**13:30–14:00**    *Business Meeting*

**2026-08-04 (continued)**

**14:00–15:30 Session O4: Oral Session 4**

Room: PAIS 290

*From Propositional to Perceptual Asymmetry: Extending FPO to Asymmetric Partial Information Dialogue*

Yifan Zhu, Kyeongmin Rim and James Pustejovsky

*Disentangling Annotator Skill from Verifier Strictness in Cross-Verified Dialogue Annotation*

Zihao Tao, John A. Prado, Ignazio Steven LaManna, Ryan Puterbaugh, Mim Datta and Julia Hirschberg

*Beyond Supervised Clarification: Input Rewriting with LLMs for Dialogue Discourse Parsing*

Yiming Liu, Ziyue Zhang, Zhichao Xu, Xin Yu, Yingheng Tang, Tianyu Jiang and Jie Cao

*Weakly Supervised Temporal Modeling of Latent Dynamics in Dyadic Conversations*

Tahiya Chowdhury

*LLMs Out-of-the-Box Do Not Generate Context-Appropriate Word Order in Russian*

Alina Shabaeva, John Frederick Bailyn and Owen Rambow

**15:30–16:00 Coffee Break**

**16:00–17:30 Session P2: Poster Session 2**

Room: PAIS main lobby

*Chronological Thinking in Full-Duplex Spoken Dialogue Language Models*

Donghang Wu, Haoyang Zhang, Chen Chen, Tianyu Zhang, Fei Tian, Xuerui Yang, Gang Yu, Hexin Liu, Nana Hou, Yuchen Hu and Eng Siong Chng

*Exploring Retrieval Augmented Generation Approaches for Natural Language Question Understanding*

Christoph Kowalski, Amelie Sophie Robrecht-Hilbig, Vincent Emmerling and Stefan Kopp

*Tree-Based Interview Topic Guidance for Collecting Target Information under Adaptive Topic continuation/switching*

Fuminori Nagasawa, Ekai Hashimoto and Shun Shiramatsu

*COCORELI: Enforcing Execution Preconditions for Reliable Collaborative Instruction Following*

Swarnadeep Bhar, Omar Naim, Eleni Metheniti, Loïc Cabannes, Bastien Navarri, Morteza Kamaladdini Ezzabady and Nicholas Asher

**2026-08-04 (continued)**

*StagePilot: Stage-Level Planning for Long-Horizon Dialogue Simulation in Cyber-grooming*

Heajun An, Qi Zhang, Minqian Liu, Xinyi Zhang, Sang Won Lee, Lifu Huang, Pamela Wisniewski and Jin-Hee Cho

*Generation of Dialogue Breakdown Repair Utterance in Non-task Oriented Conversation*

Kazuya Tsubokura, Yurie Iribe and Norihide Kitaoka

*From Sound to Symptom: Real-Time Respiratory Signal Understanding for Conversational Healthcare Agents*

Tanmay Laud, Herprit Mahal and Subhabrata Mukherjee

*Interpretable Satisfaction Modeling in Sales Dialogues: A Metrics-based Approach Validated Against Text-based Model*

Natalia Krawczyk, Alicja Kasicka, Bartosz Przybył and Justyna Gromada

*Adaptive Emotion Management in Human-robot Dialogue using Online Group Relative Policy Optimization*

Anna Manaseryan and Casey Kennington

*PersonaKit (PK): A Plug-and-Play Platform for User Testing Diverse Roles in Simulated Full-Duplex Dialogue*

Hyunbae Jeon and Jinho D. Choi

*Question Types for Knowledge Acquisition in LLM-based Dialogue Systems: Experiments with Simulated and Human Users*

Kazunori Komatani, Ryu Takeda and Mikio Nakano

*TRACER: Early Failure Detection for Task-Oriented Dialogue*

Erfan Nourbakhsh, Rocky Slavin, Ke Yang and Anthony Rios

*Relax: A Common Denominator for Connectives and Discourse Relations in French*

Fatou Sow, Yannick Toussaint and Mathieu Constant

**19:00–21:30    *Social Event***

**2026-08-05**

**08:30–09:00**    *Registration*

**09:00–10:00**    *Keynote 3: Dilek Hakkani-Tür*

**10:00–10:30**    *Coffee Break*

**10:30–12:00**    **Session O5: Oral Session 5**  
Room: PAIS 290

*MemeBuddy: Dialog-Style Audio Representations for Engaging Non-Visual Meme Experiences*

Chirag Bhansali, Vikas Ashok and Hae-Na Lee

*On the Structure of Address in Multi-Party Dialogue: From Discrete Labels to Continuous Levels*

Taiga Mori, Koji Inoue, Divesh Lala and Tatsuya Kawahara

*Bridging Talk and Thought: Understanding Dialogue Dynamics Across Collaborative Problem-Solving Contexts*

Zhengyuan Liu, Stella Yin, Min-Yen Kan and Nancy Chen

*Seeing Is Not Sharing: Some Vision-Language Models Overestimate Common Ground in Asymmetric Dialogue*

Nan Li, Albert Gatt and Massimo Poesio

**12:00–13:30**    *Lunch Break*

2026-08-05 (continued)

**13:30–15:00 Session P3: Poster Session 3**

Room: PAIS main lobby

*Evaluation of Paralinguistic-Aware Spoken Dialogue Systems using Next-Utterance Classification*

Kouki Miyazawa and Yoshinao Sato

*Better Scores, Worse Grounding: Hidden Regressions after Fine-Tuning in Dialogue Fact Verification*

Hyunkyung Park and Arkaitz Zubiaga

*LLMs and their Limited Theory of Mind: Evaluating Mental State Annotations in Situated Dialogue*

Katharine Kowalyshyn and Matthias Scheutz

*Enhancing Operational Safety via Agentic Dialogue Hazard Identification Analysis*

Sanjay Das, Ran Elgedawy, Ethan Seefried, Ryan Burchfield and Tirthankar Ghosal

*Building Task-Oriented Dialogue Systems via Instruction Guidance without Annotated Data*

Henry Gao and Jinho D. Choi

*Bounded Conversational Memory with Hybrid Retrieval and Evidence Highlighting for Multi-Session Dialogue Systems*

Manasa Mudunuri, Chandra Kiran Guntupalli, Mario Michael Kubek, Shiraj Pokharel and Murray Patterson

*Too Polite to Disagree: Understanding Sycophancy Propagation in Multi-Agent Systems*

Vira Kasprova, Amruta Parulekar, Abdulrahman AlRabah, Krishna Agaram, Ritwik Garg, Sagar Jha, Nimet Beyza Bozdog and Dilek Hakkani-Tur

*When the Database Fails: Prompting LLM Dialogue Agents for Safe Recovery in Task-Oriented Dialogue*

Mohammad Alijanpour Shalmani, Alale Rezvani Boroujeni and Jiann S. Yuan

*MTDiag: A Multi-Turn Diagnostic Dataset Towards Clinically Meaningful LLM Evaluation*

Pia Chouayfati, Alexander M. Fichtl, Miriam Anschütz, George Doumat and Georg Groh

*PERSONAJUDGE: Simulating Individual Human Preference Judgments with Evaluator-Specific Demonstration Data*

Zeyu He, Xuan Qi, Subramanian Chidambaram, Zhichao Xu, Vinayak Arannil, Lydia Chilton and Alex C. Williams

*Fractional Decay KV-Cache: Ownership-Aware Memory Management for Improved Inference Relevancy in Dialog Systems*

Sukanta Ganguly



**2026-08-05 (continued)**

*A Reinforcement Learning-Based Facilitator for Simulated Group Motivational Interviewing*

Alafate Abulimiti, Vladislav Maraev, Agnès Helme-Guizon and Catherine Pelachaud

**15:00–15:30** *Coffee Break*

**15:30–16:00** *Invited Talk: Matthew Marge (DARPA) - Frontier AI for National Security: The AI Forge Program*

**16:00–17:00** *Awards and Closing*

