

StarSEM 2026

**The 15th Joint Conference on Lexical and Computational
Semantics (*SEM 2026)**

Proceedings of the Conference (*SEM 2026)

July 3, 2026

The StarSEM organizers gratefully acknowledge the support from the following sponsors.

Sponsored by



©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 979-8-89176-413-2

Preface by the Program Chairs

We are excited to welcome you to *SEM 2026, the 15th Joint Conference on Lexical and Computational Semantics! We are pleased to present this volume containing the accepted long and short papers. *SEM 2026 is being held on July 3rd, 2026, in San Diego, California, co-located with ACL 2026. Since its first edition in 2012, *SEM has become a major venue to present recent advances in all areas of lexical and computational semantics, including semantic representations, theoretical semantics, multilingual semantics, and others. *SEM is sponsored by SIGLEX, the ACL Special Interest Group on the Lexicon.

*SEM 2026 accepted papers submitted both directly to *SEM and those already reviewed through ARR (ACL Rolling Review). We received submissions in various areas, including:

- Lexical Semantics
- Semantic Composition and Sentence-level Semantics
- Discourse, Dialogue and Generation
- Commonsense Reasoning and NLU
- Resources and Evaluation
- Theoretical and Formal Semantics
- Multilinguality
- Semantics in NLP Applications
- Psycholinguistics, Cognitive Linguistics, and Semantic Processing
- Social Biases and Ethics
- Interpretability and Explainability

We had 64 submissions this year combining both direct submissions and ARR commits. We compiled an exciting and wide-ranging program, accepting a total of 36 papers (34 long papers and 2 short papers). The submitted papers were carefully evaluated by a program committee led by 9 area chairs, who coordinated a large team of reviewers.

This year we started a new initiative to encourage reflection on the type of research question being explored both while writing a paper and in the review process (<https://starsem2026.github.io/blog>). Centering and explicitly articulating the research question helps authors frame and present their contribution more clearly. Just as importantly, it helps reviewers and Area Chairs evaluate the work within the appropriate context. For instance, a paper that centers a compelling linguistic or behavioral research question and offers meaningful new insights need not also introduce methodological novelty or rely on the latest models (including LLMs). Collecting this information in the submission form also allows *SEM to track the kinds of research questions authors pursue and how the conference’s focus evolves over time.

In the final set of accepted papers of 2026, approximately 74% focused on new findings about automatic language processing (through ML/AI and other approaches), 24% on new findings about language (linguistic phenomena, semantic patterns), and 2% on new findings about people (language use, behavior, health, ethics, etc.). The percentages were broadly similar in the submitted papers.

The program committee also selected a paper for the Best Paper award and an additional paper for Honorable mention. We are also delighted to have two excellent keynote speakers: Dallas Card (University of Michigan) and Alex Warstadt (UC San Diego).

We are honored to serve in the organizing committee for *SEM 2026, and we absolutely could not have made this happen without a huge amount of help. We are especially grateful to the 2025 Program Co-Chairs Lea Frermann and Mark Stevenson (also General Chair for 2026) for thoughtful guidance. Thanks to all the Area Chairs and reviewers for helping shape a rigorous and diverse conference program. Thanks to our Publicity Chair Hsuvas Borkakoty (IT:U Interdisciplinary Transformation University Austria) for managing the website and social media updates. Thanks to our Publication Chairs Zheng Yuan (Uni-

versity of Sheffield) and Yi (Jodie) Zhou (Cardiff University) for putting together the proceedings, and to the ACL 2026 workshop organizers for help and support with all organizational aspects of the conference. Finally, thank you to the authors and presenters for making *SEM 2026 such an engaging and exciting event!

We hope that you will find the contributions in these proceedings as engaging as we do, and we hope to see you at future iterations of *SEM!

Nedjma Ousidhoum and Saif M. Mohammad, Program Chairs

Organizing Committee

General Chair

Mark Stevenson, University of Sheffield

Program Chairs

Nedjma Ousidhoum, Cardiff University

Saif M. Mohammad, National Research Council Canada

Publication Chairs

Zheng Yuan, University of Sheffield

Yi (Jodie) Zhou, Cardiff University

Publicity Chair

Hsuvas Borkakoty, IT:U Interdisciplinary Transformation University Austria

Program Committee

Area Chairs

Valerio Basile, University of Turin
Steven Bethard, University of Arizona
Hila Gonen, University of British Columbia
Bridget T. McInnes, Virginia Commonwealth University
Siyao Peng, LMU Munich
Steven Schockaert, Cardiff University
Sowmya Vajjala, National Research Council, Canada
Genta Indra Winata, Capital One
Yi Zhou, Cardiff University

Reviewers

Timothee Mickus, University of Helsinki
Prasanth Yadla, Birla Institute of Technology and Science
Jing Yang Lee, Nanyang Technological University
Koji Mineshima, Keio University
Eugenio Martínez-Cámara, Universidad de Jaén
Nona Naderi, Université Paris-Saclay
Enrica Troiano, HK3Lab
Andrey Kutuzov, University of Oslo
Eduardo Blanco, University of Arizona
Grzegorz Kondrak, University of Alberta
Ian Porada, Otter.ai
Antonio Jimeno Yepes, Unstructured Technologies
Alda Mari, CNRS
Abhishek Bhandari, Indian Institute of Technology, Jodhpur
Valia Kordoni, Humboldt Universität Berlin
Kin Ian Lo, University College London, University of London
Paul Cook, University of New Brunswick
Halil Kilicoglu, University of Illinois at Urbana-Champaign
Manikandan Ravikiran, Thoughtworks Research and IIT Mandi
Ning Shi, University of Alberta
Ameeta Agrawal, Portland State University
Alessandra Zarcone, Technische Hochschule Augsburg
Nikolai Ilinykh, Göteborg University
Abhidip Bhattacharyya, University of Massachusetts at Amherst
Mohamad Hardyman Barawi, Universiti Malaysia Sarawak
Simona Frenda, Heriot-Watt University
Ted Pedersen, University of Minnesota - Duluth
Narjes Tahaei, Concordia University
Yoshihiko Hayashi, Waseda University
Ratna Kandala, University of Kansas
Sam Henry, Randolph-Macon College
Akash Anil, Indian Institute of Science Education and Research Bhopal
Sabine Schulte im Walde, University of Stuttgart

Michael Bloodgood, The College of New Jersey
Francis Ferraro, University of Maryland, Baltimore County
Davide Buscaldi, Ecole polytechnique and Université Paris 13
Lawrence S. Moss, Indiana University at Bloomington
Rossella Varvara, University of Turin
Timothée Bernard, Université Paris Cité
Tai Vu, OpenAI
Colin S. Gordon, Drexel University
Yisong Miao, National University of Singapore
Vethavikashini Chithrra Raghuram, Amazon
Ruchira Dhar, University of Copenhagen
Naomi Baes, University of Melbourne
Gabriella Chronis, University of Texas at Austin
Richard Moot, CNRS and LIRMM
Leonardo Ranaldi, University of Edinburgh
Kenneth Lai, Brandeis University and Mass General Brigham
Tak-Lam Wong, Douglas College
Voula Giouli, Aristotle University of Thessaloniki and ILSP - AthenaResearch Center
Jing Chen, Hong Kong Polytechnic University
Eduardo Calò, Utrecht University
Hanzi Xu, Temple University
Bonaventura Coppola, SAP Security Research
Venkata S Govindarajan, Ithaca College
Robin Cooper, University of Gothenburg
Md. Faiyaz Abdullah Sayeedi, Independent University, Bangladesh
Vered Shwartz, University of British Columbia
Joanne Boisson, Cardiff University
Parush Gera, Palo Alto Networks
Emerson Cabrera Paraiso, Pontificia Universidade Católica do Paraná
Dongqiang Yang, Shandong Jianzhu University
Dongqiang Yang, Shandong Jianzhu University
Sriharsha Hatwar, CapitalOne
Adarsa Sivaprasad, University of Aberdeen
Ratna Kandala, University of Kansas
Yuwei Yin, University of British Columbia
Maksym Taranukhin, University of British Columbia
Longfei Zuo, Technical University of Munich
Beiduo Chen, Ludwig-Maximilians-Universität München
Michael Oliverio, University of Turin
Eliana Di Palma, University of Turin
Cristina Bosco, University of Turin
Marta Marchiori Manerba, University of Turin
Sveva Silvia Pasini, University of Turin
Jonas Becker, Georg-August Universität Göttingen
Liang-Chih Yu, Yuan Ze University
Amy L. Olex, Virginia Commonwealth University
Sam Henry, Randolph-Macon College
Pratip Rana, Old Dominion University
Tianhui Zhang, University of Liverpool
Pingjun Hong, University of Vienna

Table of Contents

<i>SemToken: Semantic-Aware Tokenization for Efficient Long-Context Language Models</i> Dong Liu and Yanxuan Yu	1
<i>Syntactic Priming in Few-Shot Learning: How Demonstration Structure Shapes LLM Performance</i> Prasanth Yadla	13
<i>A Logic-Based Approach to Hallucinations in Data-to-Text NLG: Experiments with Human and LLM Annotators</i> Eduardo Calò, Saad Mahamood, Albert Gatt and Kees Van Deemter	28
<i>A framework for annotating and modelling intentions behind metaphor use</i> Gianluca Michelli, Xiaoyu Tong and Ekaterina Shutova	63
<i>ReFRAME or Remain: Unsupervised Lexical Semantic Change Detection with Frame Semantics</i> Bach Phan Tat, Kris Heylen, Stefano De Pascale, Dirk Geeraerts and Dirk Speelman	83
<i>Finding Sense in Nonsense with Generated Contexts: Perspectives from Humans and Language Models</i> Katrina Olsen and Sebastian Padó	98
<i>Uncovering Ideological Bias in RAG with Lexical Multidimensional Analysis: A Case Study on COVID-19</i> Elmira Salari, Maria Claudia Nunes Delfino, Hazem Amamou, José Victor de Souza, Shruti Kshirsagar, Alan Davoust and Anderson Avila	111
<i>Evaluating Adjective-Noun Compositionality in LLMs: Functional vs Representational Perspectives</i> Ruchira Dhar, Qiwei Peng and Anders Sjøgaard	125
<i>Lexical Availability and Human Distributional Agreement in GPT-4o’s Color Naming</i> Anna Feldman and Jing Peng	136
<i>The LSCD Benchmark: a Testbed for Diachronic Word Meaning Tasks</i> Dominik Schlechtweg, Sachin Yadav, Jonas Kuhn and Nikolay Arefyev	148
<i>From Latents to Labels: Zero-Shot Named Entity Recognition using Sparse Autoencoder Features</i> Nakanyseth Vuth, Gilles Sérasset and Didier Schwab	164
<i>Can You Be More Explicit? A Task and Dataset on Explicitations of Implicit Meaning</i> Laura Zeidler and Michael Roth	178
<i>Compositional Meaning Representations in LLMs: a Critical Review of Probing Studies</i> Rémy Marro	198
<i>*-PLUIE: Personalisable metric with Llm Used for Improved Evaluation</i> Quentin Lemesle, Leane Jourdan, Daisy Munson, Pierre Alain, Jonathan Chevelu, Arnaud Delhay and Damien Lolive	211
<i>Under the Surface: Probing Tamil Paraphrase Intelligence</i> Viswadarshan R R, Dr. J. Felicia Lilian and Mahalakshmi S	244
<i>Annotating Indian Regional Biases using Large Language Models: Evaluation and Analysis</i> Debasmita Panda, Akash Anil and Neelesh Kumar Shukla	255

<i>The Tatoxa System for Text Detoxification in Low-Resource Languages: The Case of Tatar</i>	
Ilseyar Alimova, Bogdan Monogov, Artyom Mazur, Daniil Antonov, Vsevolod Karimov, Vitaliy Egorov, Bulat Khakimov and Alexander Panchenko	264
<i>Towards Sense-level Bilingual Dictionary Induction</i>	
Lydia Körber, Katja Markert and Wei Zhao	275
<i>Beyond Accuracy: Interpreting Topic Representation in Suicide Ideation Detection Models</i>	
Hamideh Ghanadian, Isar Nejadgholi and Hussein Al Osman	290
<i>Understanding the Linguistic Cues Behind Stance Detection</i>	
Parush Gera and Tempestt Neal	302
<i>Diagnosing Generalization in Open-Source LLMs for Stance Detection</i>	
Parush Gera and Tempestt Neal	317
<i>Can LLMs Solve My Grandma’s Riddle? Evaluating Multilingual Large Language Models on Reasoning Traditional Bangla Tricky Riddles</i>	
Nurul Labib Sayeedi, Md. Faiyaz Abdullah Sayeedi, Khushnur Binte Jahangir, Swakkhar Shatabda and Sarah Masud Preum	331
<i>Undocumented Immigrants “!= Illegal Aliens”: Decomposing the Conceptual and Narrative Landscapes of Partisan Immigration Terms</i>	
Yejin Cho, Gabriella Chronis, Nitin Sudarsanam, Kevin Barcenas-Martinez and Katrin Erk	348
<i>Can a Remedy Find a Researcher? Exploring the Development of Semantic Knowledge in Italian BabyLMs</i>	
Alice Suozzi, Luca Capone, Gianluca Leboni and Alessandro Lenci	366
<i>Frame In, Frame Out: Measuring Framing Bias in LLM-Generated News Summaries</i>	
Valeria Pastorino and Nafise Sadat Moosavi	378
<i>Mitigating Language Bias in Multilingual Sentence Embeddings for Cross-Lingual Similarity Estimation</i>	
Kanade Nonomura, Keita Fukushima, Risa Kondo and Tomoyuki Kajiwara	385
<i>Entropy-aware Masking for Masked Language Modeling</i>	
Gokul Srinivasagan, Kai Hartung and Munir Georges	395
<i>Do Factual Recall Mechanisms Carry over from Text to Speech in Multimodal Language Models?</i>	
Luca Modica, Filip Landin, Mehrdad Farahani, Livia Qian, Gabriel Skantze and Richard Johansson	401
<i>Interactive Agents: Simulating Counselor-Client Psychological Counseling via Role-Playing LLM-to-LLM Interactions</i>	
Huachuan Qiu and Zhenzhong Lan	410
<i>ZIP: Quantifying Which Words Matter in Zero-Shot Instructional Prompts</i>	
Nikta Gohari Sadr, Sangmitra Madhusudan, Arash Asgari, Hassan Sajjad, Laleh Seyyed-Kalantari and Ali Emami	428
<i>The Frequency Confound in Language-Model Surprisal and Metaphor Novelty</i>	
Omar Momen and Sina Zarrieß	454
<i>One Word Is Not Enough: Simple Prompts Improve Word Embeddings</i>	
Rajeev Ranjan	464

<i>HistoryBankQA: Multilingual Temporal Question Answering on Historical Events</i> Biswadip Mandal, Anant Khandelwal and Manish Gupta	474
<i>EvalMORAAL: Interpretable Chain-of-Thought and LLM-as-Judge Evaluation for Moral Alignment in Large Language Models</i> Hadi Mohammadi, Anastasia Giachanou and Robert A. Bagheri	497
<i>Toxic Subword Pruning for Dialogue Response Generation on Large Language Models</i> Hongyuan Adam Lu and Wai Lam	516
<i>Dictionary Insertion Prompting for Multilingual Reasoning on Multilingual Large Language Models</i> Hongyuan Lu, Z.L. L and Wai Lam	529