

COLING 2025

**Proceedings of the
1st Workshop on Computational Humor (CHum)**

Editors

Christian F. Hempelmann, Julia Rayz,
Tiansi Dong, and Tristan Miller

January 19, 2025

©2025 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

ISBN 979-8-89176-204-6

Preface

The impressive recent progress in generative AI has opened up new frontiers for complex tasks that can be thought of as essentially human, including the production or interpretation of humor. For its part, humor research in general has become a mature, interdisciplinary field, both in theoretical advances across disciplines such as psychology, linguistics, and sociology, and in its breadth of purview and empirical support. While much existing research raises the question of whether modern language models can meaningfully contribute to theoretical linguistics, this First Workshop on Computational Humor (CHum 2025) proposes a parallel inquiry: despite their primary focus on engineering objectives, should these models be evaluated for their relevance to humor theory? Specifically, we set out to examine whether they can effectively model humor in alignment with established theoretical frameworks. As such, the workshop aims to foster further work on modeling the processes of humor with current methods in computational linguistics and natural language processing, against the theoretical backdrop of humor research and with reference to relevant corpora of textual, visual, and multimodal materials.

Being more dependent on context and inference than straightforward forms of communication, humor poses particular challenges for probabilistic approaches. For these and other reasons, it has long served as a special field of application for computational linguistics – understanding humor has often been referred to as “AI complete” and as such considered to be a potential test case for Artificial General Intelligence. Humor is approached presumably for this very reason whenever new paradigms arise in NLP, which currently center around large language models (LLMs). However, the study of humor is inherently multidisciplinary, and computational humor systems should therefore integrate insights from multiple disciplines in order to perform effectively. With this in mind, the workshop’s keynotes bring together researchers from theoretical linguistics, psychology, and computer science, all of whom have expertise in humor theories. Among the questions raised in the keynotes are how elements of humor theories can be incorporated into LLM-generated text, and whether LLMs can be used to test humor theories.

Although we have chosen to name our event the “First” workshop for computational humor, we must acknowledge the groundwork that has been laid by a number of previous gatherings. These include the Twelfth Twente Workshop on Language Technology joint with the International Workshop on Computational Humour (1996), the Twentieth Twente Workshop on Language Technology, titled The April Fool’s Day Workshop On Computational Humour (2002), the 3rd International Workshop on Computational Humor (2012), the 2012 AAI Fall Symposium Artificial Intelligence of Humor (2012), Dagstuhl Seminar 21362 on Structure and Learning (2021), and the annual panels on Humor and Artificial Intelligence at the International Society of Humor Studies Conferences (2018–). These meetings, sometimes limited to invited contributions, aimed to explore diverse goals, from assessing how modeling humor could contribute to modeling intelligence, to the design of intelligent systems capable of understanding the (theoretical) mechanisms of jokes and collaboratively contributing to their generation. Like us, many of our invited speakers have participated in these previous meetings, making them well positioned to provide insights into the progress and evolution of the field.

CHum 2025 received a total of 28 submissions, which is many more than we could accept for presentation. In the end, we selected eleven papers that we felt most closely aligned with the workshop’s interdisciplinary aims. Our selection includes work on advancing the state of the art in humor research by applying current computational models and techniques, as well as more task-focused NLP papers that apply techniques informed by existing humor research.

We would like to thank everyone who submitted a paper to the workshop, as well as the members of our Program Committee for their timely and insightful reviews.

Keynotes

How Many Stochastic Parrots Does it Take to Change a Lightbulb?

Salvatore Attardo, East Texas A&M University

Generally speaking the discussion of humor and AI has centered on questions such as “Does AI have a sense of humor?” or “Does AI understand humor and/or can it explain it?” or “Can AI produce humor?” More pragmatic projects tend to focus on recognition: “Can AI identify humor/irony/puns/satire?” I find these questions interesting, but up to a point. First, they seem all tinged, more or less consciously, to what we could call the “anthropomorphic AI ideology”: the belief that AIs match/reflect/approximate human intelligence and that humor is a folk-metric of performance, a sort of Humorous Turing test: if AIs understand/produce humor then they are truly “human” (for some senses of “be” and “human” to be specified). The character of Data in *Star Trek: The Next Generation* is the standard bearer for this set of beliefs. Second, the idea, to which I shamefully contributed, that humor is “AI complete” and hence if an AI “masters” humor it should thereby be an AGI (Artificial General Intelligence). Third, the current wave of LLMs are essentially black boxes and it is unclear to what extent a black box model counts as an explanation. Having managed to alienate the entire audience of the gathering, I will then proceed to argue that LLMs provide us with an excellent implementation of Trier’s (1931) lexical field theory and more broadly of the semantic network postulated by Fillmore, Raskin, Eco and many others. As such they make some aspects of theories of humor testable in a new empirical way.

Unlocking the Punchline: Navigating Humor Computation From Understanding to Generation and Beyond

Liang Yang, Dalian University of Technology

Significant progress has been made in the field of humor computation, including recognition and generation. Despite its promising developments, several challenges still remain in the era of LLMs. In this talk, I will present our recent efforts to address some of them. (1) For understanding, I will discuss the understanding mechanism of large models for humor. (2) For generation, I will introduce how large models generate specific types of humor by integrating humor theory. (3) Finally, I will share our future exploration directions, such as evaluation of sense of humor.

The Psychology of Computational Humor

Willibald Ruch, University of Zurich

This keynote will highlight what computational humor needs from research into the psychology of humor, what psychology can gain from requirements and results in computational humor, and where this underappreciated multidisciplinary symbiosis has (not) gone and needs to develop.

Get With The Program! From Talking the Talk to Walking the Walk in Computational Humour

Tony Veale, University College Dublin

LLM-based agents such as ChatGPT exhibit an amiable if superficial wit, and show a skilled ear for mimicry that lets them capture the cadences and attitudes of famous comics. Their auto-regressive generation of continuations to a given prompt makes them especially adept at using the “yes, and . . .” principle of improv comedy to wring laughs from silly premises. But jokes of the short and snappy variety that are crafted to be retold in many different contexts require an LLM to do more than talk like a comedian. The words and attitudes must be appropriate to a joke, and the text must actually work like a joke too, with an effective logical mechanism that snaps shut on an audience like a mouse in a trap. Although LLMs can do a surprisingly good job of analyzing and explaining jokes, they do a rather poor job of turning this reactive appreciation into a proactive generative ability. In this talk I consider how we might tackle this *understanding vs. generating* gap.

Organizing Committee

Christian F. Hempelmann, East Texas A&M University
Julia Rayz, Purdue University
Tiansi Dong, Fraunhofer IAIS and University of Cambridge
Tristan Miller, University of Manitoba

Program Committee

Adam Jatowt, University of Innsbruck
Alessandro Valitutti, Phedes Lab
Anna Palmann, University of Amsterdam
Anne-Gwenn Bosser, École Nationale d'Ingénieurs de Brest
Bob Mankoff, CartoonCollections.com
Drew Gorenz, University of Southern California
Elena Mikhalkova, European University at Saint Petersburg
Heather Knight, Oregon State University
Joe Toplyn, Twenty Lane Media
Juliette Love, Google DeepMind
Kory Mathewson, Google DeepMind
Kristian Kersting, TU Darmstadt
Larry Lefkowitz, Working Knowledge LLC
Max Petrenko, Amazon
Nathaniel Laywine, York University
Ori Amir, University of Texas Permian Basin
Pelin Cunningham-Erdogdu, University of Texas at Austin
Piotr Mirowski, Google DeepMind
Rada Mihalcea, University of Michigan
Rey Alejandro Gonzalez, Purdue University
Sophie Jentzsch, German Aerospace Center (DLR)
Tatiana Ringenber, Purdue University
Yi Zhang, Purdue University

Table of Contents

<i>The Exception of Humor: Iconicity, Phonemic Surprisal, Memory Recall, and Emotional Associations</i> Alexander Kilpatrick and Maria Flaksman	1
<i>Text Is Not All You Need: Multimodal Prompting Helps LLMs Understand Humor</i> Ashwin Baluja	9
<i>Rule-based Approaches to the Automatic Generation of Puns Based on Given Names in French</i> Mathieu Dehouck and Marine Delaborde	18
<i>Homophonic Pun Generation in Code Mixed Hindi English</i> Yash Raj Sarrof	23
<i>Bridging Laughter Across Languages: Generation of Hindi-English Code-mixed Puns</i> Likhith Asapu, Prashant Kodali, Ashna Dua, Kapil Rajesh Kavitha and Manish Shrivastava	32
<i>Testing Humor Theory Using Word and Sentence Embeddings</i> Stephen Skalicky and Salvatore Attardo	58
<i>Pragmatic Metacognitive Prompting Improves LLM Performance on Sarcasm Detection</i> Joshua Lee, Wyatt Fong, Alexander Le, Sur Shah, Kevin Han and Kevin Zhu	63
<i>Can AI Make Us Laugh? Comparing Jokes Generated by Witscript and a Human Expert</i> Joe Toplyn and Ori Amir	71
<i>Evaluating Human Perception and Bias in AI-Generated Humor</i> Narendra Nath Joshi	79
<i>The Theater Stage as Laboratory: Review of Real-Time Comedy LLM Systems for Live Performance</i> Piotr Mirowski, Kory Mathewson and Boyd Branch	88
<i>The Algorithm is the Message: Computing as a Humor-Generating Mode</i> Vittorio Marone	96

Workshop Program

08:45–09:00 *Opening remarks*
Christian F. Hempelmann, Julia Rayz, Tiansi Dong and Tristan Miller

Keynote 1

09:00–09:45 *How Many Stochastic Parrots Does it Take to Change a Lightbulb?*
Salvatore Attardo, East Texas A&M University

Paper session 1: Multimodal Pathways to Meaning and Memory

09:45–10:05 *The Exception of Humor: Iconicity, Phonemic Surprisal, Memory Recall, and Emotional Associations*
Alexander Kilpatrick and Maria Flaksman

10:05–10:25 *Text Is Not All You Need: Multimodal Prompting Helps LLMs Understand Humor*
Ashwin Baluja

Keynote 2

11:00–11:45 *Unlocking the Punchline: Navigating Humor Computation From Understanding to Generation and Beyond*
Liang Yang, Dalian University of Technology

Paper session 2: Algorithms and Frameworks for Pun Generation

11:45–12:05 *Rule-based Approaches to the Automatic Generation of Puns Based on Given Names in French*
Mathieu Dehouck and Marine Delaborde

12:05–12:25 *Homophonic Pun Generation in Code Mixed Hindi English*
Yash Raj Sarrof

12:25–12:45 *Bridging Laughter Across Languages: Generation of Hindi-English Code-mixed Puns*
Likhith Asapu, Prashant Kodali, Ashna Dua, Kapil Rajesh Kavitha and Manish Shrivastava

Keynote 3

13:30–14:15 *The Psychology of Computational Humor*
Willibald Ruch, University of Zurich

Paper session 3: Explorations in Semantics and Pragmatics

14:15–14:35 *Testing Humor Theory Using Word and Sentence Embeddings*
Stephen Skalicky and Salvatore Attardo

14:55–15:40 *Pragmatic Metacognitive Prompting Improves LLM Performance on Sarcasm Detection*
Joshua Lee, Wyatt Fong, Alexander Le, Sur Shah, Kevin Han and Kevin Zhu

Keynote 4

14:55–15:40 *Get With The Program! From Talking the Talk to Walking the Walk in Computational Humour*
Tony Veale, University College Dublin

Paper session 4: Evaluating and Experiencing Generated Humor

16:10–16:30 *Can AI Make Us Laugh? Comparing Jokes Generated by Witscript and a Human Expert*
Joe Toplyn and Ori Amir

16:30–16:50 *Evaluating Human Perception and Bias in AI-Generated Humor*
Narendra Nath Joshi

16:50–17:10 *The Theater Stage as Laboratory: Review of Real-Time Comedy LLM Systems for Live Performance*
Piotr Mirowski, Kory Mathewson and Boyd Branch

17:10–17:30 *The Algorithm is the Message: Computing as a Humor-Generating Mode*
Vittorio Marone

17:30–17:40 *Closing remarks*
Christian F. Hempelmann, Julia Rayz, Tiansi Dong and Tristan Miller

