

CDL 2026

**The 1st Workshop on Computational Developmental  
Linguistics (CDL)**

**Proceedings of the Workshop**

July 3, 2026

©2026 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

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

ISBN 979-8-89176-428-6

# Introduction

We are pleased to welcome you to the First Workshop on Computational Developmental Linguistics (CDL), held in conjunction with ACL 2026 in San Diego.

Computational developmental linguistics brings together questions about language acquisition, learning dynamics, adaptation, and change over time across human and artificial learners. The workshop is intended as a venue for researchers in machine learning, computational linguistics, cognitive science, and developmental psychology who are interested in studying how linguistic knowledge emerges, evolves, and adapts.

**Background** The first workshop on computational developmental linguistics (CDL) invites interdisciplinary contributions broadly in the topic of computational developmental linguistics. By computational developmental linguistics, we broadly refer to the studies of computational modeling of language acquisition and change over time, encompassing both individual learners, including humans or machines, throughout their lifetime and across populations. Particularly, we welcome submissions that study the developmental trajectories through which learners master phonology, lexicon, and syntax, as well as the idiolectal shifts that arise as speakers adapt to new communities, domains, or interlocutors. With the advent of neural language models (LMs), we now have large-scale systems that can be examined for learning dynamics, representational change, knowledge tracing, and language adaptation. Beyond static text corpora, interactive and multimodal training regimes make it possible to reverse-engineer developmental conditions, explore resource-efficient learning, and measure parallels and divergences between human and machine language acquisition. This rapidly growing area not only advances our understanding of how linguistic capabilities emerge and evolve in artificial systems, but also provides new potential for generating hypotheses about human language learning.

**Scope and Goal** This workshop aims to bridge the conversation between modern machine learning and developmental linguistics. We hope to draw inspiration from both fields, identifying similarities and differences in the (im)plausibilities of comparing LMs and human language learning, to motivate better LM engineering and more rigorous computational modeling of human language acquisition. Possible comparisons include (1) pretraining dynamics vs. first language acquisition: relating the data and scaling behavior of large models to the developmental timelines, stages, and milestones documented in human language learning; (2) continual learning and post-training vs. idiolect change and second language acquisition: exploring how ongoing adaptation, fine-tuning, or domain shifts in large models mirror semantic drift and individual language change in human speakers. By bringing together researchers from machine learning, computational linguistics, cognitive science, and developmental psychology, we aim to foster cross-disciplinary dialogue and establish frameworks for building computational models and performing computational analyses of language development, enabling scientifically rigorous comparisons between humans and machines without overly anthropomorphizing LMs.

**Topics of Interest** The scope and topics include, but are not limited to:

- Computational models for developmental linguistics. Models and formalisms for simulating first and second language acquisition, including data constraints, model architectures, training objectives, and frameworks for modeling idiolect change and semantic drift.
- Learning dynamics in pretraining and post-training of LMs. Behavioral and mechanistic interpretations of how linguistic and cognitive competence emerge during pretraining, and how it evolves through post-training (e.g., finetuning, alignment, and domain adaptation). In addition to text-only LMs, we encourage approaches that integrate multiple modalities and interactive learning, as these more closely mirror human language acquisition processes.

- Comparisons of developmental linguistics in humans and machines. Comparing the processes, constraints, and outcomes of language acquisition in humans and artificial learners; examining the possibilities, limitations, and methodological challenges of such comparisons; relating model learning trajectories to developmental timelines and milestones in humans; and conducting these analyses in a scientifically rigorous manner without over-anthropomorphizing LMs.
- Knowledge tracing and developmental diagnostics. Developing tools and methods to track the acquisition, transformation, and retention of linguistic knowledge in LMs, including phase transition detection, benchmarks, and probes of representational change.
- Applications in language education and clinical NLP. Applying insights from computational developmental linguistics to improve LM engineering, language tutoring systems, child-directed AI interaction, adaptive educational technology, and computational models of neurological communication disorders.

For this first edition of the workshop, we were particularly excited to see contributions spanning language modeling, speech and multimodal learning, child-directed data, interaction, semantic change, and the analysis of developmental trajectories in both human and machine language systems. The workshop also features invited talks and a panel discussion intended to support broad interdisciplinary exchange.

We thank the authors for their submissions and the reviewers for their time, care, and expertise. We are also grateful to the invited speakers, panelists, and organizers whose work made this workshop possible. We hope that CDL helps establish a durable meeting place for this emerging research area and supports future work at the intersection of developmental linguistics and modern machine learning.

Martin Ziqiao Ma, Emmy Liu, Jing Liu, Tyler A. Chang, Abdellah Fourtassi, Alex Warstadt, Michael Hahn, Weiwei Sun, and Freda Shi

# Organizing Committee

## Organizing Committee

Martin Ziqiao Ma, University of Michigan

Emmy Liu, Carnegie Mellon University

Jing Liu, École Normale Supérieure

Tyler A. Chang, Google DeepMind

Abdellah Fourtassi, Aix-Marseille University

Alex Warstadt, UC San Diego

Michael Hahn, Saarland University

Weiwei Sun, University of Cambridge

Freda Shi, University of Waterloo / Vector Institute

# Program Committee

## Area Chairs

Tyler A. Chang, Google DeepMind  
Abdellah Fourtassi, Aix Marseille University  
Michael Hahn, Saarland University  
Emmy Liu, School of Computer Science, Carnegie Mellon University  
Ziqiao Ma, Thinking Machines Lab  
Freda Shi, Vector Institute and University of Waterloo

## Reviewers

Bastian Bunzeck

Sarah C. Creel

Abdellah Fourtassi

Jugal Gajjar, Sadaf Ghaffari, Prachi Goyal

Nadia Ghezaiel Hammouda

Parin Rajesh Jhaveri

Shunsuke Kando, Garry Kuwanto

Siddharth Lall, Dan Le, Xue Li

Fred Mailhot, Akshata Kishore Moharir

Ruoxi Ning, Sergiu Nisioi

Udita Patel, Van-Thuy Phi

Mengyang Qiu

Suchir Salhan, Shivani Shekhar, Gyu-Ho Shin, Siyuan Song, Xiaonan Song, Dewang Sultania

Rohith Uppala

Tai Vu

Yixuan Wang, Ahmed Wez, Daniel Wurgaft

Aditya Yadavalli

Chen Zhang, Jiayi Zhang

## Table of Contents

<i>Linguistics Theory Meets LLM: Code-Switched Text Generation via Equivalence Constrained Large Language Models</i> Garry Kuwanto, Chaitanya Agarwal, Genta Indra Winata and Derry Tanti Wijaya .....	1
<i>Do Structural Priors Help Neural Language Models Learn Grammar? Evidence from Child-Scale Data</i> Jon-Paul Cacioli .....	15
<i>Fine-tuned speech representations track spoken language convergence to adult models in infants and children who are deaf/hard-of-hearing</i> Landon Choy, Ali Sartaz Khan, Sonia Patrizi, Daisy S. Ye, Julianna Gross and Margaret Cychosz	27
<i>Do Language Models Show Structural Priming Across Different Domains?</i> So Young Lee, Russell Scheinberg and Ameeta Agrawal .....	37
<i>Do large language models and humans follow similar learning stages? Assessing GPT-2's order of Swedish grammar acquisition within the Processability Theory framework</i> Stella Lundqvist, Murathan Kurfali and Johan Sjons .....	52
<i>On the Learnability of Syntax from Raw Speech with Autoregressive Predictive Coding</i> Shunsuke Kando and Yusuke Miyao .....	77
<i>Modeling Writing Development as Coordinated Change Across Linguistic and Semantic Dimensions</i> Michelle Banawan, Andrew Potter, Tracy Arner and Danielle S McNamara .....	83
<i>L1 Influence in L2 Language Models: A Human-centric Approach</i> Laura Barbenel, Lily Goulder, Aoife O'Driscoll, Suchir Salhan, Catherine Arnett, Andrew Caines and Paula Buttery .....	92
<i>A Scalable Tool for Measuring Manner and Result Verbs in Developmental Language Research</i> Divyesh Pratap Singh, Dakshesh Gusain, Federica Bulgarelli, Alison Eisel Hendricks, John Beavers, Nathan M. Beers and Ifeoma Nwogu .....	117
<i>Making Synthetic Questions More Child-Directed: Prompting and Sampling Effects</i> Whitney Poh, Michael Tombolini and Libby Barak .....	129