

NAACL 2025

**The 22nd SIGMORPHON workshop on Computational
Morphology, Phonology, and Phonetics**

Proceedings of the Workshop

May 3, 2025

©2025 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 979-8-89176-231-2

Organizing Committee

Co-Chair

Garrett Nicolai, University of British Columbia
Eleanor Chodroff, University of York
Çagri Çöltekin, University of Tübingen
Fred Mailhot, Dialpad, Inc.

SIGMORPHON Officers

President: Garrett Nicolai, University of British Columbia
At Large: Eleanor Chodroff, University of York
At Large: Çagri Çöltekin, University of Tübingen
At Large: Fred Mailhot, Dialpad, Inc.

Program Committee

Michael Ginn, University of Colorado
Kenneth Steimel, Cisco Systems Incorporated
Sandra Kübler, Indiana University
Changbing Yang, University of British Columbia
Sarah Moeller, University of Florida
Rob Malouf, San Diego State University
Gregory Kobele, Leipzig University

Keynote Talk

Invited Talk

Kyle Gorman

City University of New York

2025-05-03 09:30:00 –

Abstract: The past and future of grapheme-to-phoneme conversion

I will describe three strands of recent work on grapheme-to-phoneme conversion (G2P). After introducing this task, I will briefly review some lessons learned from the SIGMORPHON 2020 and 2021 shared tasks on G2P. Then, I will focus on low-resource scenarios, describing in-progress work on unsupervised G2P using a decipherment framework and novel algorithms. Finally, I will turn to high-resource scenarios, describing in-progress work on what I am calling “grapheme-and-phoneme-to-phoneme conversion”, in which “side pronunciations” from other dialects and/or transcription systems are provided as additional inputs to the G2P system to improve robustness and generalization.

Bio: Kyle Gorman is an associate professor at the Graduate Center, City University of New York, where he directs the master’s program in computational linguistics. He is also a software engineer at Google Research. He is president and cofounder of the ACL Special Interest Group on Writing Systems and Written Language (SIGWrit). He is the coauthor of Finite-State Text Processing (Morgan & Claypool, 2021) and papers in computational linguistics and phonological theory.

Keynote Talk

Panel Discussion

Gasper Begus, Kyle Gorman, Connor Mayer, Kristine Yu

University of California Berkeley, City University of New York, University of Massachusetts Amherst,
University of California Irvine

2025-05-03 13:30:00 –

Bio: Gašper Begus is an Associate Professor at the University of California, Berkeley, where he bridges the fields of linguistics, artificial intelligence, and cognitive science. As the Principal Investigator of the Berkeley Speech and Computation lab, Dr. Beguš leads efforts to develop the most realistic models of human language learning—creating an artificial baby language learner. Serving as the Linguistics Lead at Project CETI, a non-profit dedicated to listening to sperm whales, he researches one of the most intriguing communication systems in the animal kingdom.

Kyle Gorman is an associate professor at the Graduate Center, City University of New York, where he directs the master’s program in computational linguistics. He is also a software engineer at Google Research. He is president and cofounder of the ACL Special Interest Group on Writing Systems and Written Language (SIGWrit). He is the coauthor of Finite-State Text Processing (Morgan & Claypool, 2021) and papers in computational linguistics and phonological theory.

Kristine Yu is an Associate Professor in the Department of Linguistics at the University of Massachusetts Amherst. Her main area of research falls in phonetics and phonology, particularly in prosody and its interfaces. She works in these areas by integrating methods and perspectives from linguistic fieldwork, laboratory experimentation, and computational methods from formal language theory to machine learning.

Connor Mayer is a phonologist and phonetician who studies phonological and phonetic learning, as well as the structure of the phonological knowledge speakers acquire. His research relies on a combination of computational and experimental approaches, with phonological theory serving as the nexus between them. His current research activities fall into three broad categories: phonotactic learning, the phonetics and phonology of Uyghur (Turkic: China), and the influence of speech motor control and biomechanics on phonological systems.

Table of Contents

<i>Prompt and circumstance":A word-by-word LLM prompting approach to interlinear glossing for low-resource languages</i>	
Micha Elsner and David Liu	1
<i>West Germanic noun-noun compounds and the morphology-syntax trade-off</i>	
Pablo Mosteiro, Damián Blasi and Denis Paperno	15
<i>The Impact of Dialect Variation on Robust Automatic Speech Recognition for Catalan</i>	
Zachary Hopton and Eleanor Chodroff	23
<i>Probing Neural Network Generalization using Default Patterns</i>	
Brandon Prickett, Tianyi Nyu and Katya Pertsova	34

Program

Saturday, May 03, 2025

09:25 - 09:30 *Opening Remarks*

09:30 - 10:30 *Invited Talk: Kyle Gorman*

10:30 - 11:00 *Break*

11:00 - 12:00 *Paper Session*

Probing Neural Network Generalization using Default Patterns

Brandon Prickett, Tianyi Nyu and Katya Pertsova

The Impact of Dialect Variation on Robust Automatic Speech Recognition for Catalan

Zachary Hopton and Eleanor Chodroff

West Germanic noun-noun compounds and the morphology-syntax trade-off

Pablo Mosteiro, Damián Blasi and Denis Paperno

Prompt and circumstance": A word-by-word LLM prompting approach to interlinear glossing for low-resource languages

Micha Elsner and David Liu

12:00 - 13:30 *Lunch*

13:30 - 14:30 *Panel: Gasper Begus, Kyle Gorman, Connor Mayer, Kristine Yu*

14:30 - 15:30 *Town Hall*

15:30 - 16:00 *Break*