

# **20th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology (SIGMORPHON 2023)**

Toronto, Canada  
14 July 2023

ISBN: 978-1-7138-8180-3

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2023) by the Association for Computational Linguistics  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Translating a low-resource language using GPT-3 and a human-readable dictionary</i> Micha Elsner and Jordan Needle . . . . .	1
<i>Evaluating Cross Lingual Transfer for Morphological Analysis: a Case Study of Indian Languages</i> Siddhesh Pawar, Pushpak Bhattacharyya and Partha Talukdar . . . . .	14
<i>Joint Learning Model for Low-Resource Agglutinative Language Morphological Tagging</i> Gulinigeer Abudouwaili, Kahaerjiang Abiderexiti, Nian Yi and Aishan Wumaier . . . . .	27
<i>Revisiting and Amending Central Kurdish Data on UniMorph 4.0</i> Sina Ahmadi and Aso Mahmudi . . . . .	38
<i>Investigating Phoneme Similarity with Artificially Accented Speech</i> Margot Masson and Julie Carson-berndsen . . . . .	49
<i>Generalized Glossing Guidelines: An Explicit, Human- and Machine-Readable, Item-and-Process Convention for Morphological Annotation</i> David R. Mortensen, Ela Gulsen, Taiqi He, Nathaniel Robinson, Jonathan Amith, Lindia Tjuatja and Lori Levin . . . . .	58
<i>Jambu: A historical linguistic database for South Asian languages</i> Aryaman Arora, Adam Farris, Samopriya Basu and Suresh Kolichala . . . . .	68
<i>Lightweight morpheme labeling in context: Using structured linguistic representations to support linguistic analysis for the language documentation context</i> Bhargav Shandilya and Alexis Palmer . . . . .	78
<i>Improving Automated Prediction of English Lexical Blends Through the Use of Observable Linguistic Features</i> Jarem Saunders . . . . .	93
<i>Colexifications for Bootstrapping Cross-lingual Datasets: The Case of Phonology, Concreteness, and Affectiveness</i> Yiyi Chen and Johannes Bjerva . . . . .	98
<i>Character alignment methods for dialect-to-standard normalization</i> Yves Scherrer . . . . .	110
<i>SIGMORPHON–UniMorph 2023 Shared Task 0: Typologically Diverse Morphological Inflection</i> Omer Goldman, Khuyagbaatar Batsuren, Salam Khalifa, Aryaman Arora, Garrett Nicolai, Reut Tsarfaty and Ekaterina Vylomova . . . . .	117
<i>SIGMORPHON–UniMorph 2023 Shared Task 0, Part 2: Cognitively Plausible Morphophonological Generalization in Korean</i> Canaan Breiss and Jinyoung Jo . . . . .	126
<i>Morphological reinflection with weighted finite-state transducers</i> Alice Kwak, Michael Hammond and Cheyenne Wing . . . . .	132
<i>Linear Discriminative Learning: a competitive non-neural baseline for morphological inflection</i> Cheonkam Jeong, Dominic Schmitz, Akhilesh Kakolu Ramarao, Anna Stein and Kevin Tang . . . . .	138
<i>Tü-CL at SIGMORPHON 2023: Straight-Through Gradient Estimation for Hard Attention</i> Leander Gırrbach . . . . .	151

<i>The BGU-MeLeL System for the SIGMORPHON 2023 Shared Task on Morphological Inflection</i> Gal Astrach and Yuval Pinter .....	166
<i>Tü-CL at SIGMORPHON 2023: Straight-Through Gradient Estimation for Hard Attention</i> Leander Gırrbach .....	171
<i>Findings of the SIGMORPHON 2023 Shared Task on Interlinear Glossing</i> Michael Ginn, Sarah Moeller, Alexis Palmer, Anna Stacey, Garrett Nicolai, Mans Hulden and Miikka Silfverberg .....	186
<i>LISN @ SIGMORPHON 2023 Shared Task on Interlinear Glossing</i> Shu Okabe and François Yvon .....	202
<i>SigMoreFun Submission to the SIGMORPHON Shared Task on Interlinear Glossing</i> Taqi He, Lindia Tjuatja, Nathaniel Robinson, Shinji Watanabe, David R. Mortensen, Graham Neubig and Lori Levin .....	209
<i>An Ensembled Encoder-Decoder System for Interlinear Glossed Text</i> Edith Coates .....	217
<i>Glossy Bytes: Neural Glossing using Subword Encoding</i> Ziggy Cross, Michelle Yun, Ananya Apparaju, Jata MacCabe, Garrett Nicolai and Miikka Silfver- berg .....	222
<i>The SIGMORPHON 2022 Shared Task on Cross-lingual and Low-Resource Grapheme-to-Phoneme Conversion</i> Arya D. McCarthy, Jackson L. Lee, Alexandra DeLucia, Travis Bartley, Milind Agarwal, Lucas F.E. Ashby, Luca Del Signore, Cameron Gibson, Reuben Raff and Winston Wu .....	230
<i>SIGMORPHON 2022 Shared Task on Grapheme-to-Phoneme Conversion Submission Description: Se- quence Labelling for G2P</i> Leander Gırrbach .....	239
<i>Low-resource grapheme-to-phoneme mapping with phonetically-conditioned transfer</i> Michael Hammond .....	245
<i>A future for universal grapheme-phoneme transduction modeling with neuralized finite-state transdu- cers</i> Chu-Cheng Lin Lin .....	249
<i>Fine-tuning mSLAM for the SIGMORPHON 2022 Shared Task on Grapheme-to-Phoneme Conversion</i> Dan Garrette .....	250