

7th Workshop on Research in Computational Typology and Multilingual NLP (SIGTYP 2025)

Vienna, Austria
1 August 2025

ISBN: 979-8-3313-2418-6

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© (2025) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

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

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
Web: www.proceedings.com

Table of Contents

<i>InstructionCP: A Simple yet Effective Approach for Transferring Large Language Models to Target Languages</i>	
Kuang-Ming Chen, Jenq-Neng Hwang and Hung-yi Lee	1
<i>Analyzing the Linguistic Priors of Language Models with Synthetic Languages</i>	
Alessio Tosolini and Terra Blevins	7
<i>Unstable Grounds for Beautiful Trees? Testing the Robustness of Concept Translations in the Compilation of Multilingual Wordlists</i>	
David Snee, Luca Ciucci, Arne Rubehn, Kellen Parker Van Dam and Johann-Mattis List	16
<i>Annotating and Inferring Compositional Structures in Numeral Systems Across Languages</i>	
Arne Rubehn, Christoph Rzymiski, Luca Ciucci, Katja Bocklage, Alžběta Kučerová, David Snee, Abishek Stephen, Kellen Parker Van Dam and Johann-Mattis List	29
<i>Beyond the Data: The Impact of Annotation Inconsistencies in UD Treebanks on Typological Universals and Complexity Assessment</i>	
Antoni Brosa Rodríguez and M. Dolores Jiménez López	43
<i>Beyond cognacy</i>	
Gerhard Jäger	52
<i>SenWiCh: Sense-Annotation of Low-Resource Languages for WiC using Hybrid Methods</i>	
Roksana Goworek, Harpal Singh Karicut, Hamza Shezad, Nijaguna Darshana, Abhishek Mane, Syam Bondada, Raghav Sikka, Ulvi Mammadov, Rauf Allahverdiyev, Sriram Satkirti Purighella, Paridhi Gupta, Muhinyia Ndegwa, Bao Khanh Tran and Haim Dubossarsky	61
<i>XCOMPS: A Multilingual Benchmark of Conceptual Minimal Pairs</i>	
Linyang He, Ercong Nie, Sukru Samet Dindar, Arsalan Firoozi, Van Nguyen, Corentin Puffay, Riki Shimizu, Haotian Ye, Jonathan Brennan, Helmut Schmid, Hinrich Schuetze and Nima Mesgarani	75
<i>Tone in Perspective: A Computational Typological Analysis of Tone Function in ASR</i>	
SiYu Liang and Gina-Anne Levow	82
<i>A discovery procedure for synlexification patterns in the world's languages</i>	
Hannah S. Rognan and Barend Beekhuizen	93
<i>Construction-Based Reduction of Translationese for Low-Resource Languages: A Pilot Study on Bavarian</i>	
Peiqin Lin, Marion Thaler, daniela.goschala@campus.lmu.de daniela.goschala@campus.lmu.de, Amir Hossein Kargaran, Yihong Liu, Andre Martins and Hinrich Schuetze	114
<i>High-Dimensional Interlingual Representations of Large Language Models</i>	
Bryan Wilie, Samuel Cahyawijaya, Junxian He and Pascale Fung	122
<i>Domain Meets Typology: Predicting Verb-Final Order from Universal Dependencies for Financial and Blockchain NLP</i>	
Zichao Li and Zong Ke	156
<i>Token-level semantic typology without a massively parallel corpus</i>	
Barend Beekhuizen	165
<i>Are Translated Texts Useful for Gradient Word Order Extraction?</i>	
Amanda Kann	177