

# **Ancient Language Processing Workshop (RANLP-ALP 2023)**

Varna, Bulgaria  
8 September 2023

ISBN: 978-1-7138-9044-7

**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>Training and Evaluation of Named Entity Recognition Models for Classical Latin</i> Marijke Beersmans, Evelien de Graaf, Tim Van de Cruys and Margherita Fantoli .....	1
<i>Sentence Embedding Models for Ancient Greek Using Multilingual Knowledge Distillation</i> Kevin Krahn, Derrick Tate and Andrew C. Lamicela .....	13
<i>A Transformer-based parser for Syriac morphology</i> Martijn Naaijer, Constantijn Sikkels, Mathias Coeckelbergs, Jisk Attema and Willem Th. Van Peursen .....	23
<i>Graecia capta ferum victorem cepit. Detecting Latin Allusions to Ancient Greek Literature</i> Frederick Riemenschneider and Anette Frank .....	30
<i>Larth: Dataset and Machine Translation for Etruscan</i> Gianluca Vico and Gerasimos Spanakis .....	39
<i>Evaluation of Distributional Semantic Models of Ancient Greek: Preliminary Results and a Road Map for Future Work</i> Silvia Stopponi, Nilo Pedrazzini, Saskia Peels, Barbara McGillivray and Malvina Nissim .....	49
<i>Latin Morphology through the Centuries: Ensuring Consistency for Better Language Processing</i> Federica Gamba and Daniel Zeman .....	59
<i>Cross-Lingual Constituency Parsing for Middle High German: A Delexicalized Approach</i> Ercong Nie, Helmut Schmid and Hinrich Schütze .....	68
<i>Can Large Language Model Comprehend Ancient Chinese? A Preliminary Test on ACLUE</i> Yixuan Zhang and Haonan Li .....	80
<i>Unveiling Emotional Landscapes in Plautus and Terentius Comedies: A Computational Approach for Qualitative Analysis</i> Davide Picca and Caroline Richard .....	88
<i>Morphological and Semantic Evaluation of Ancient Chinese Machine Translation</i> Kai Jin, Dan Zhao and Wuying Liu .....	96
<i>A tailored Handwritten-Text-Recognition System for Medieval Latin</i> Philipp Koch, Gily Vera Nuñez, Esteban Garces Arias, Christian Heumann, Matthias Schöffel, Alexander Häberlin and Matthias Assenmacher .....	103
<i>Evaluating Existing Lemmatizers on Unedited Byzantine Greek Poetry</i> Colin Swaelens, Ilse De Vos and Els Lefever .....	111
<i>Vector Based Stylistic Analysis on Ancient Chinese Books: Take the Three Commentaries on the Spring and Autumn Annals as an Example</i> Yue Qi, Liu Liu, Bin Li and Dongbo Wang .....	117
<i>A Joint Model of Automatic Word Segmentation and Part-Of-Speech Tagging for Ancient Classical Texts Based on Radicals</i> Bolin Chang, Yiguo Yuan, Bin Li, Zhixing Xu, Minxuan Feng and Dongbo Wang .....	122

<i>Introducing an Open Source Library for Sumerian Text Analysis</i> Hansel Guzman-Soto and Yudong Liu .....	133
<i>Coding Design of Oracle Bone Inscriptions Input Method Based on "ZhongHuaZiKu" Database</i> Dongxin Hu .....	138
<i>Word Sense Disambiguation for Ancient Greek: Sourcing a training corpus through translation alignment</i> Alek Keersmaekers, Wouter Mercelis and Toon Van Hal .....	148
<i>Enhancing State-of-the-Art NLP Models for Classical Arabic</i> Tariq Yousef, Lisa Mischer, Hamid Reza Hakimi and Maxim Romanov .....	160
<i>Logion: Machine-Learning Based Detection and Correction of Textual Errors in Greek Philology</i> Charlie Cowen-Breen, Creston Brooks, Barbara Graziosi and Johannes Haubold .....	170
<i>Classical Philology in the Time of AI: Exploring the Potential of Parallel Corpora in Ancient Languages</i> Tariq Yousef, Chiara Palladino and Farnoosh Shamsian .....	179
<i>Using Word Embeddings for Identifying Emotions Relating to the Body in a Neo-Assyrian Corpus</i> Ellie Bennett and Aleksí Sahala .....	193
<i>A Neural Pipeline for POS-tagging and Lemmatizing Cuneiform Languages</i> Aleksi Sahala and Krister Lindén .....	203
<i>Tibetan Dependency Parsing with Graph Convolutional Neural Networks</i> Bo An .....	213
<i>On the Development of Interlinearized Ancient Literature of Ethnic Minorities: A Case Study of the Interlinearization of Ancient Written Tibetan Literature</i> Congjun Long and Bo An .....	222