

# **16th Workshop on Innovative Use of NLP for Building Educational Applications 2021**

Online  
20 April 2021

ISBN: 978-1-7138-2957-7

**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© (2021) by the Association for Computational Linguistics  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

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>Negation Scope Resolution for Chinese as a Second Language</i> Mengyu Zhang, Weiqi Wang, Shuqiao Sun and Weiwei Sun .....	1
<i>Text Simplification by Tagging</i> Kostiantyn Omelianchuk, Vipul Raheja and Oleksandr Skurzhanskyi .....	11
<i>Employing distributional semantics to organize task-focused vocabulary learning</i> Haemant Santhi Ponnusamy and Detmar Meurers .....	26
<i>Synthetic Data Generation for Grammatical Error Correction with Tagged Corruption Models</i> Felix Stahlberg and Shankar Kumar .....	37
<i>Broad Linguistic Complexity Analysis for Greek Readability Classification</i> Savvas Chatzipanagiotidis, Maria Giagkou and Detmar Meurers .....	48
<i>Character Set Construction for Chinese Language Learning</i> Chak Yan Yeung and John Lee .....	59
<i>Identifying negative language transfer in learner errors using POS information</i> Leticia Farias Wanderley and Carrie Demmans Epp .....	64
<i>Document-level grammatical error correction</i> Zheng Yuan and Christopher Bryant .....	75
<i>Essay Quality Signals as Weak Supervision for Source-based Essay Scoring</i> Haoran Zhang and Diane Litman .....	85
<i>Parsing Argumentative Structure in English-as-Foreign-Language Essays</i> Jan Wira Gotama Putra, Simone Teufel and Takenobu Tokunaga .....	97
<i>Training and Domain Adaptation for Supervised Text Segmentation</i> Goran Glavaš, Ananya Ganesh and Swapna Somasundaran .....	110
<i>Data Strategies for Low-Resource Grammatical Error Correction</i> Simon Flachs, Felix Stahlberg and Shankar Kumar .....	117
<i>Towards a Data Analytics Pipeline for the Visualisation of Complexity Metrics in L2 writings</i> Thomas Gaillat, Anas Knefati and Antoine Lafontaine .....	123
<i>Estonian as a Second Language Teacher's Tools</i> Tiiu Üksik, Jelena Kallas, Kristina Koppel, Katrin Tsepelina and Raili Pool .....	130
<i>Assessing Grammatical Correctness in Language Learning</i> Anisia Katinskaia and Roman Yangarber .....	135
<i>On the application of Transformers for estimating the difficulty of Multiple-Choice Questions from text</i> Luca Benedetto, Giovanni Aradelli, Paolo Cremonesi, Andrea Cappelli, Andrea Giussani and Roberto Turrin .....	147
<i>Automatically Generating Cause-and-Effect Questions from Passages</i> Katherine Stasaski, Manav Rathod, Tony Tu, Yunfang Xiao and Marti A. Hearst .....	158

<i>Interventions Recommendation: Professionals' Observations Analysis in Special Needs Education</i> Javier Muñoz and Felipe Bravo-Marquez.....	171
<i>C-Test Collector: A Proficiency Testing Application to Collect Training Data for C-Tests</i> Christian Haring, Rene Lehmann, Andrea Horbach and Torsten Zesch .....	180
<i>Virtual Pre-Service Teacher Assessment and Feedback via Conversational Agents</i> Debajyoti Datta, Maria Phillips, James P. Bywater, Jennifer Chiu, Ginger S. Watson, Laura Barnes and Donald Brown .....	185
<i>Automated Classification of Written Proficiency Levels on the CEFR-Scale through Complexity Contours and RNNs</i> Elma Kerz, Daniel Wiechmann, Yu Qiao, Emma Tseng and Marcus Ströbel .....	199
<i>"Sharks are not the threat humans are": Argument Component Segmentation in School Student Essays</i> Tariq Alhindi and Debanjan Ghosh .....	210
<i>Using Linguistic Features to Predict the Response Process Complexity Associated with Answering Clinical MCQs</i> Victoria Yaneva, Daniel Jurich, Le An Ha and Peter Baldwin .....	223