

NAACL HLT 2016

**The Eleventh Workshop
on Innovative Use of NLP for
Building Educational Applications**

Proceedings of the Workshop

June 16, 2016
San Diego, California, USA

Gold Sponsors



Silver Sponsors



©2016 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-941643-83-9

Introduction

We are excited to be holding the 11th edition of the BEA workshop. Since starting in 1997, the BEA workshop, now one of the largest workshops at NAACL/ACL, has become one of the leading venues for publishing innovative work that uses NLP to develop educational applications. The consistent interest and growth of the workshop has clear ties to challenges in education, especially with regard to supporting literacy. The research presented at the workshop illustrates advances in the technology, and the maturity of the NLP/education field that are responses to those challenges with capabilities that support instructor practices and learner needs. NLP capabilities now support an array of learning domains, including writing, speaking, reading, and mathematics. In the writing and speech domains, automated writing evaluation (AWE) and speech assessment applications, respectively, are commercially deployed in high-stakes assessment and instructional settings, including Massive Open Online Courses (MOOCs). We also see widely-used commercial applications for plagiarism detection and peer review. There has been a renewed interest in spoken dialog and multi-modal systems for instruction and assessment as well as feedback. We are also seeing explosive growth of mobile applications for game-based applications for instruction and assessment. The current educational and assessment landscape, especially in the United States, continues to foster a strong interest and high demand that pushes the state-of-the-art in AWE capabilities to expand the analysis of written responses to writing genres other than those traditionally found in standardized assessments, especially writing tasks requiring use of sources and argumentative discourse.

The use of NLP in educational applications has gained visibility outside of the NLP community. First, the Hewlett Foundation reached out to public and private sectors and sponsored two competitions: one for automated essay scoring, and another for scoring of short answer, subject-matter-based response items. The motivation driving these competitions was to engage the larger scientific community in this enterprise. MOOCs are now beginning to incorporate AWE systems to manage the thousands of constructed-response assignments collected during a single MOOC course. Learning@Scale is another venue that discusses NLP research in education. The Speech and Language Technology in Education (SLaTE), now in its seventh year, promotes the use of speech and language technology for educational purposes. Another breakthrough for educational applications within the CL community is the presence of a number of shared-task competitions over the last three years. There have been four shared tasks on grammatical error correction with the last two held at CoNLL (2013 and 2014). In 2014 alone, there were four shared tasks for NLP and Education-related areas. We are pleased to announce a unique shared task at BEA this year: Automated Evaluation of Scientific Writing.

As a community, we continue to improve existing capabilities, and to identify and generate innovative ways to use NLP in applications for writing, reading, speaking, critical thinking, curriculum development, and assessment. Steady growth in the development of NLP-based applications for education has prompted an increased number of workshops, typically focusing on one specific subfield. In this workshop, we present papers from the following subfields: tools for automated scoring of text and speech, automated test-item generation, curriculum development, collaborative problem solving, content evaluation in text, dialogue and intelligent tutoring, evaluation of genres beyond essays, feedback studies, and grammatical error detection.

This year we received a record 46 submissions, and accepted 8 papers as oral presentations and 20 as poster presentation and/or demos, for an overall acceptance rate of 61%. Each paper was reviewed by three members of the Program Committee who were believed to be most appropriate for each paper. We continue to have a very strong policy to deal with conflicts of interest. First, we made a concerted effort to not assign papers to reviewers to evaluate if the paper had an author from their institution. Second, with respect to the organizing committee, authors of papers for which there was a conflict of interest recused themselves from the discussions.

While the field is growing, we do recognize that there is a core group of institutions and researchers who work in this area. With a higher acceptance rate, we were able to include papers from a wider variety of topics and institutions. The papers accepted were selected on the basis of several factors, including the relevance to a core educational problem space, the novelty of the approach or domain, and the strength of the research. The accepted papers were highly diverse – an indicator of the growing variety of foci in this field. We continue to believe that the workshop framework designed to introduce work in progress and new ideas needs to be revived, and we hope that we have achieved this with the breadth and variety of research accepted for this workshop, a brief description of which is presented below:

For *automated writing evaluation*, Meyer & Koch investigate how users of intelligent writing assistance tools deal with correct, incorrect, and incomplete feedback; Rei & Cummins investigate the task of assessing sentence-level prompt relevance in learner essays; Cummins et al focus on determining the topical relevance of L2 essays to the prompt; Loukina & Cahill investigate how well systems developed for automated evaluation of written responses perform when applied to spoken responses; Beigman Klebanov et al address the problem of quantifying the overall extent to which a test-taker's essay deals with the topic it is assigned; King & Dickinson investigate questions of how to reason about learner meaning in cases where the set of correct meanings is never entirely complete, specifically for the case of picture description tasks; Madnani et al present preliminary work on automatically scoring tests of proficiency in music instruction; Rahimi & Litman automatically extract and investigate the usefulness of topical components for scoring the Evidence dimension of an analytical writing in response to text assessment; Ledbetter & Dickinson describe the development of a morphological analyzer for learner Hungarian, outlining extensions to a resource-light system that can be developed by different types of experts.

For *short-answer scoring*, Horbach & Palmer explore the suitability of active learning for automatic short-answer assessment on the ASAP corpus; Banjade et al present a corpus that contains student answers annotated for their correctness in context, in addition to a baseline for predicting the correctness label; and Rudzewitz explores the practical usefulness of the combination of features from three different fields – short answer scoring, authorship attribution, and plagiarism detection – for two tasks: semantic learner language classification, and plagiarism detection for evaluating short answers.

For *grammar and spelling error detection*, Madnani et al discuss a classifier approach that yields higher

precision and a language modeling approach that provides better recall; Beinborn et al discuss a model that can predict spelling difficulty with a high accuracy, and provide a thorough error analysis that takes the L1 into account and provides insights into cross-lingual transfer effects; Napoles et al estimate the deterioration of NLP processing given an estimate of the amount and nature of grammatical errors in a text; and, Yuan et al develop a supervised ranking model to re-rank candidates generated from an SMT-based grammatical error correction system.

For *text difficulty and curriculum development*, Xia et al address the task of readability assessment for texts aimed at L2 learners; Reynolds investigates Russian second language readability assessment using a machine-learning approach with a range of lexical, morphological, syntactic, and discourse features; Chen & Meurers study the frequency of a word in common language use, and systematically explore how such a word-level feature is best used to characterize the reading levels of texts; Yoon et al present an automated method for estimating the difficulty of spoken texts for use in generating items that assess non-native learners' listening proficiency; Milli & Hearst explore the automated augmentation of a popular online learning resource – Khan Academy video modules – with relevant reference chapters from open access textbooks; and Chinkina & Meurers present an IR system for text selection that identifies the grammatical constructions spelled out in the official English language curriculum of schools in Baden-Württemberg (Germany) and re-ranks the search results based on the selected (de)prioritization of grammatical forms.

For *item generation*, Hill & Simha propose a method to automatically generate multiple-choice fill-in-the-blank exercises from existing text passages that challenge a reader's comprehension skills and contextual awareness; Wojatzki et al present the concept of bundled gap filling, along with an efficient computational model for automatically generating unambiguous gap bundle exercises, and a disambiguation measure for guiding the construction of the exercises and validating their level of ambiguity; and Pilán explores the factors influencing the dependence of single sentences on their larger textual context in order to automatically identify candidate sentences for language learning exercises from corpora which are presentable in isolation.

For *collaborative problem solving*, Flor et al present a novel situational task that integrates collaborative problem solving behavior with testing in a science domain.

For *accessibility*, Martinez-Santiago et al discuss computer-designed tools in order to help people with Autism Spectrum Disorder to palliate or overcome such verbal limitations.

As noted earlier, this year we are excited to host the first Shared Task in *Automated Evaluation of Scientific Writing* (<http://textmining.lt/aesw/index.html>). The task involves automatically predicting whether sentences found in scientific language are in need of editing. Six teams competed and their system description papers are found in these proceedings and are presented as posters in conjunction with the BEA11 poster session. A summary report of the shared task (Daudaravicius et al) is also found in the proceedings and will be presented orally.

We wish to thank everyone who showed interest and submitted a paper, all of the authors for their

contributions, the members of the Program Committee for their thoughtful reviews, and everyone who attended this workshop. We would especially like to thank our sponsors; at the Gold Level: American Institutes for Research (AIR), Cambridge Assessment, Educational Testing Service, Grammarly, Pacific Metrics and Turnitin / Lightside, and at the Silver Level: Cognii and iLexIR. Their contributions allow us to subsidize students at the workshop dinner, and make workshop t-shirts! We would like to thank Joya Tetreault for creating the t-shirt design (again!).

Joel Tetreault, Yahoo

Jill Burstein, Educational Testing Service

Claudia Leacock, Grammarly

Helen Yannakoudakis, University of Cambridge

Organizers:

Joel Tetreault, Yahoo Labs
Jill Burstein, Educational Testing Services
Claudia Leacock, Grammarly
Helen Yannakoudakis, University of Cambridge

Program Committee:

Laura Allen, Arizona State University
Rafael Banchs, I2R
Timo Baumann, Universität Hamburg
Lee Becker, Hapara
Beata Beigman Klebanov, Educational Testing Service
Lisa Beinborn, Technische Universität Darmstadt
Kay Berkling, Cooperative State University Karlsruhe
Suma Bhat, University of Illinois, Urbana-Champaign
Serge Bibauw, Université Catholique de Louvain
David Bloom, Pacific Metrics
Chris Brew, Thomson Reuters
Ted Briscoe, University of Cambridge
Chris Brockett, Microsoft Research
Julian Brooke, University of Melbourne
Aoife Cahill, Educational Testing Service
Lei Chen, Educational Testing Service
Min Chi, North Carolina State University
Martin Chodorow, Hunter College and the Graduate Center, CUNY
Mark Core, University of Southern California
Scott Crossley, Georgia State University
Luis Fernando D'Haro, Human Language Technology - Institute for Infocomm Research
Daniel Dahlmeier, SAP
Barbara Di Eugenio, University of Illinois Chicago
Markus Dickinson, Indiana University
Yo Ehara, Tokyo Metropolitan University
Keelan Evanini, Educational Testing Service
Mariano Felice, University of Cambridge
Michael Flor, Educational Testing Service
Thomas François, Université Catholique de Louvain
Michael Gamon, Microsoft Research
Binyam Gebrekidan Gebre, Max Planck Computing and Data Facility
Kallirroi Georgila, University of Southern California
Dan Goldwasser, Purdue University
Cyril Goutte, National Research Council Canada

Iryna Gurevych, Technische Universität Darmstadt
Na-Rae Han, University of Pittsburgh
Andrea Horbach, Saarland University
Chung-Chi Huang, National Institutes of Health
Radu Tudor Ionescu, University of Bucharest
Ross Israel, Factual
Fazel Keshtkar, Southeast Missouri State University
Ekaterina Kochmar, University of Cambridge
Mamoru Komachi, Tokyo Metropolitan University
Bob Krovetz, Lexical Research
Lun-Wei Ku, Academia Sinica
Kristopher Kyle, Georgia State University
John Lee, City University of Hong Kong
Ben Leong, Educational Testing Service
James Lester, North Carolina State University
Diane Litman, University of Pittsburgh
Annie Louis, University of Essex
Anastassia Loukina, Educational Testing Service
Xiaofei Lu, Pennsylvania State University
Wencan Luo, University of Pittsburgh
Nitin Madnani, Educational Testing Service
Shervin Malmasi, Macquarie University
Montse Maritxalar, University of the Basque Country
Julie Medero, Harvey Mudd College
Detmar Meurers, Universität Tübingen
Lisa Michaud, Aspect Software
Rada Mihalcea, University of Michigan
Michael Mohler, Language Computer Corp.
Smaranda Muresan, Columbia University
Courtney Napoles, Johns Hopkins University
Hwee Tou Ng, National University of Singapore
Vincent Ng, University of Texas, Dallas
Huy Nguyen, University of Pittsburgh
Rodney Nielsen, University of North Texas
Nobal Niraula, The University of Memphis
Simon Ostermann, Saarland University
Alexis Palmer, Heidelberg University
Ted Pedersen, University of Minnesota, Duluth
Ildikó Pilán, University of Gothenburg
Zahra Rahimi, University of Pittsburgh
Lakshmi Ramachandran, Pearson
Arti Ramesh, University of Maryland, College Park
Marek Rei, University of Cambridge
Robert Reynolds, University of Tromsø
Brian Riordan, Educational Testing Service
Mark Rosenstein, Pearson
Mihai Rotaru, Textkernel

Alla Rozovskaya, Virginia Tech
C. Anton Rytting, University of Maryland, College Park
Keisuke Sakaguchi, Johns Hopkins University
Mathias Schulze, University of Waterloo
Swapna Somasundaran, Educational Testing Service
Helmer Strik, Centre for Language Studies, Centre for Language and Speech Technology, Rad-
boud University, Nijmegen
David Suendermann-Oeft, Educational Testing Service
Sowmya Vajjala, Iowa State University
Giulia Venturi, Institute of Computational Linguistics "Antonio Zampolli" (ILC-CNR)
Elena Volodina, University of Gothenburg
Carl Vogel, Trinity College
Xinhao Wang, Educational Testing Service
Michael White, Department of Linguistics, The Ohio State University
David Wible, National Central University
Alistair Willis, The Open University, UK
Magdalena Wolska, Universität Tübingen
Peter Wood, University of Saskatchewan
Huichao Xue, Google
Helen Yannakoudakis, University of Cambridge
Marcos Zampieri, Saarland University
Klaus Zechner, Educational Testing Service
Torsten Zesch, University of Duisburg-Essen
Fan Zhang, University of Pittsburgh
Xiaodan Zhu, National Research Council Canada

Table of Contents

<i>The Effect of Multiple Grammatical Errors on Processing Non-Native Writing</i> Courtney Napoles, Aoife Cahill and Nitin Madnani	1
<i>Text Readability Assessment for Second Language Learners</i> Menglin Xia, Ekaterina Kochmar and Ted Briscoe	12
<i>Automatic Generation of Context-Based Fill-in-the-Blank Exercises Using Co-occurrence Likelihoods and Google n-grams</i> Jennifer Hill and Rahul Simha	23
<i>Automated classification of collaborative problem solving interactions in simulated science tasks</i> Michael Flor, Su-Youn Yoon, Jiangang Hao, Lei Liu and Alina von Davier	31
<i>Computer-assisted stylistic revision with incomplete and noisy feedback. A pilot study</i> Christian M. Meyer and Johann Frerik Koch	42
<i>A Report on the Automatic Evaluation of Scientific Writing Shared Task</i> Vidas Daudaravicius, Rafael E. Banchs, Elena Volodina and Courtney Napoles	53
<i>Topicality-Based Indices for Essay Scoring</i> Beata Beigman Klebanov, Michael Flor and Binod Gyawali	63
<i>Predicting the Spelling Difficulty of Words for Language Learners</i> Lisa Beinborn, Torsten Zesch and Iryna Gurevych	73
<i>Characterizing Text Difficulty with Word Frequencies</i> Xiaobin Chen and Detmar Meurers	84
<i>Unsupervised Modeling of Topical Relevance in L2 Learner Text</i> Ronan Cummins, Helen Yannakoudakis and Ted Briscoe	95
<i>UW-Stanford System Description for AESW 2016 Shared Task on Grammatical Error Detection</i> Dan Flickinger, Michael Goodman and Woodley Packard	105
<i>Shallow Semantic Reasoning from an Incomplete Gold Standard for Learner Language</i> Levi King and Markus Dickinson	112
<i>The NTNU-YZU System in the AESW Shared Task: Automated Evaluation of Scientific Writing Using a Convolutional Neural Network</i> Lung-Hao Lee, Bo-Lin Lin, Liang-Chih Yu and Yuen-Hsien Tseng	122
<i>Automated scoring across different modalities</i> Anastassia Loukina and Aoife Cahill	130
<i>Model Combination for Correcting Preposition Selection Errors</i> Nitin Madnani, Michael Heilman and Aoife Cahill	136

<i>Pictogrammar: an AAC device based on a semantic grammar</i>	
Fernando Martínez-Santiago, Miguel Ángel García Cumbreiras, Arturo Montejo Ráez and Manuel Carlos Díaz Galiano	142
<i>Detecting Context Dependence in Exercise Item Candidates Selected from Corpora</i>	
Ildikó Pilán	151
<i>Feature-Rich Error Detection in Scientific Writing Using Logistic Regression</i>	
Madeline Remse, Mohsen Mesgar and Michael Strube	162
<i>Bundled Gap Filling: A New Paradigm for Unambiguous Cloze Exercises</i>	
Michael Wojatzki, Oren Melamud and Torsten Zesch	172
<i>Evaluation Dataset (DT-Grade) and Word Weighting Approach towards Constructed Short Answers Assessment in Tutorial Dialogue Context</i>	
Rajendra Banjade, Nabin Maharjan, Nobal Bikram Niraula, Dipesh Gautam, Borhan Samei and Vasile Rus	182
<i>Linguistically Aware Information Retrieval: Providing Input Enrichment for Second Language Learners</i>	
Maria Chinkina and Detmar Meurers	188
<i>Enhancing STEM Motivation through Personal and Communal Values: NLP for Assessment of Utility Value in Student Writing</i>	
Beata Beigman Klebanov, Jill Burstein, Judith Harackiewicz, Stacy Priniski and Matthew Mulholland	199
<i>Cost-Effectiveness in Building a Low-Resource Morphological Analyzer for Learner Language</i>	
Scott Ledbetter and Markus Dickinson	206
<i>Automatically Scoring Tests of Proficiency in Music Instruction</i>	
Nitin Madnani, Aoife Cahill and Brian Riordan	217
<i>Combined Tree Kernel-based classifiers for Assessing Quality of Scientific Text</i>	
Liliana Mamani Sanchez and Hector-Hugo Franco-Penya	223
<i>Augmenting Course Material with Open Access Textbooks</i>	
Smitha Milli and Marti A. Hearst	229
<i>Exploring the Intersection of Short Answer Assessment, Authorship Attribution, and Plagiarism Detection</i>	
Björn Rudzewitz	235
<i>Sentence-Level Grammatical Error Identification as Sequence-to-Sequence Correction</i>	
Allen Schmaltz, Yoon Kim, Alexander M. Rush and Stuart Shieber	242
<i>Combining Off-the-shelf Grammar and Spelling Tools for the Automatic Evaluation of Scientific Writing (AESW) Shared Task 2016</i>	
René Witte and Bahar Sateli	252

<i>Candidate re-ranking for SMT-based grammatical error correction</i>	
Zheng Yuan, Ted Briscoe and Mariano Felice	256
<i>Spoken Text Difficulty Estimation Using Linguistic Features</i>	
Su-Youn Yoon, Yeonsuk Cho and Diane Napolitano	267
<i>Automatically Extracting Topical Components for a Response-to-Text Writing Assessment</i>	
Zahra Rahimi and Diane Litman	277
<i>Sentence Similarity Measures for Fine-Grained Estimation of Topical Relevance in Learner Essays</i>	
Marek Rei and Ronan Cummins	283
<i>Insights from Russian second language readability classification: complexity-dependent training requirements, and feature evaluation of multiple categories</i>	
Robert Reynolds	289
<i>Investigating Active Learning for Short-Answer Scoring</i>	
Andrea Horbach and Alexis Palmer	301

Conference Program

Thursday June 16, 2016

08:45–09:00 *Load Oral Presentations*

09:00–09:15 *Opening Remarks*

09:15–09:40 *The Effect of Multiple Grammatical Errors on Processing Non-Native Writing*
Courtney Napoles, Aoife Cahill and Nitin Madnani

09:40–10:05 *Text Readability Assessment for Second Language Learners*
Menglin Xia, Ekaterina Kochmar and Ted Briscoe

10:05–10:30 *Automatic Generation of Context-Based Fill-in-the-Blank Exercises Using Co-occurrence Likelihoods and Google n-grams*
Jennifer Hill and Rahul Simha

10:30–11:00 *Break*

11:00–11:25 *Automated classification of collaborative problem solving interactions in simulated science tasks*
Michael Flor, Su-Youn Yoon, Jiangang Hao, Lei Liu and Alina von Davier

11:25–11:50 *Computer-assisted stylistic revision with incomplete and noisy feedback. A pilot study*
Christian M. Meyer and Johann Frerik Koch

11:50–12:15 *A Report on the Automatic Evaluation of Scientific Writing Shared Task*
Vidas Daudaravicius, Rafael E. Banchs, Elena Volodina and Courtney Napoles

12:25–02:00 *Lunch*

02:00–02:45 *Poster and Demo Session A*

Topicality-Based Indices for Essay Scoring
Beata Beigman Klebanov, Michael Flor and Binod Gyawali

Predicting the Spelling Difficulty of Words for Language Learners
Lisa Beinborn, Torsten Zesch and Iryna Gurevych

Thursday June 16, 2016 (continued)

Characterizing Text Difficulty with Word Frequencies

Xiaobin Chen and Detmar Meurers

Unsupervised Modeling of Topical Relevance in L2 Learner Text

Ronan Cummins, Helen Yannakoudakis and Ted Briscoe

UW-Stanford System Description for AESW 2016 Shared Task on Grammatical Error Detection

Dan Flickinger, Michael Goodman and Woodley Packard

Shallow Semantic Reasoning from an Incomplete Gold Standard for Learner Language

Levi King and Markus Dickinson

The NTNU-YZU System in the AESW Shared Task: Automated Evaluation of Scientific Writing Using a Convolutional Neural Network

Lung-Hao Lee, Bo-Lin Lin, Liang-Chih Yu and Yuen-Hsien Tseng

Automated scoring across different modalities

Anastassia Loukina and Aoife Cahill

Model Combination for Correcting Preposition Selection Errors

Nitin Madhani, Michael Heilman and Aoife Cahill

Pictogrammar: an AAC device based on a semantic grammar

Fernando Martínez-Santiago, Miguel Ángel García Cumbreiras, Arturo Montejo Ráez and Manuel Carlos Díaz Galiano

Detecting Context Dependence in Exercise Item Candidates Selected from Corpora

Ildikó Pilán

Feature-Rich Error Detection in Scientific Writing Using Logistic Regression

Madeline Remse, Mohsen Mesgar and Michael Strube

Bundled Gap Filling: A New Paradigm for Unambiguous Cloze Exercises

Michael Wojatzki, Oren Melamud and Torsten Zesch

02:45–03:30 *Poster and Demo Session B*

Evaluation Dataset (DT-Grade) and Word Weighting Approach towards Constructed Short Answers Assessment in Tutorial Dialogue Context

Rajendra Banjade, Nabin Maharjan, Nobal Bikram Niraula, Dipesh Gautam, Borhan Samei and Vasile Rus

Thursday June 16, 2016 (continued)

Linguistically Aware Information Retrieval: Providing Input Enrichment for Second Language Learners

Maria Chinkina and Detmar Meurers

Enhancing STEM Motivation through Personal and Communal Values: NLP for Assessment of Utility Value in Student Writing

Beata Beigman Klebanov, Jill Burstein, Judith Harackiewicz, Stacy Priniski and Matthew Mulholland

Cost-Effectiveness in Building a Low-Resource Morphological Analyzer for Learner Language

Scott Ledbetter and Markus Dickinson

Automatically Scoring Tests of Proficiency in Music Instruction

Nitin Madnani, Aoife Cahill and Brian Riordan

Combined Tree Kernel-based classifiers for Assessing Quality of Scientific Text

Liliana Mamani Sanchez and Hector-Hugo Franco-Penya

Augmenting Course Material with Open Access Textbooks

Smitha Milli and Marti A. Hearst

Exploring the Intersection of Short Answer Assessment, Authorship Attribution, and Plagiarism Detection

Björn Rudzewitz

Sentence-Level Grammatical Error Identification as Sequence-to-Sequence Correction

Allen Schmaltz, Yoon Kim, Alexander M. Rush and Stuart Shieber

Combining Off-the-shelf Grammar and Spelling Tools for the Automatic Evaluation of Scientific Writing (AESW) Shared Task 2016

René Witte and Bahar Sateli

Candidate re-ranking for SMT-based grammatical error correction

Zheng Yuan, Ted Briscoe and Mariano Felice

Spoken Text Difficulty Estimation Using Linguistic Features

Su-Youn Yoon, Yeonsuk Cho and Diane Napolitano

Automatically Extracting Topical Components for a Response-to-Text Writing Assessment

Zahra Rahimi and Diane Litman

03:30–04:00 Break

Thursday June 16, 2016 (continued)

04:00–04:20 *Sentence Similarity Measures for Fine-Grained Estimation of Topical Relevance in Learner Essays*

Marek Rei and Ronan Cummins

04:20–04:45 *Insights from Russian second language readability classification: complexity-dependent training requirements, and feature evaluation of multiple categories*

Robert Reynolds

04:45–05:10 *Investigating Active Learning for Short-Answer Scoring*

Andrea Horbach and Alexis Palmer

05:10–05:25 *Closing Remarks*

