Eleventh Workshop on Innovative Use of NLP for Building Educational Applications 2016

Held at the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2016)

San Diego, California, USA
16 June 2016

ISBN: 978-1-5108-2525-3
Table of Contents

The Effect of Multiple Grammatical Errors on Processing Non-Native Writing
Courtney Napoles, Aoife Cahill and Nitin Madnani ........................................ 1

Text Readability Assessment for Second Language Learners
Menglin Xia, Ekaterina Kochmar and Ted Briscoe ....................................... 12

Automatic Generation of Context-Based Fill-in-the-Blank Exercises Using Co-occurrence Likelihoods and Google n-grams
Jennifer Hill and Rahul Simha .......................................................... 23

Automated classification of collaborative problem solving interactions in simulated science tasks
Michael Flor, Su-Youn Yoon, Jiangang Hao, Lei Liu and Alina von Davier ................. 31

Computer-assisted stylistic revision with incomplete and noisy feedback. A pilot study
Christian M. Meyer and Johann Frerik Koch ............................................. 42

A Report on the Automatic Evaluation of Scientific Writing Shared Task
Vidas Daudaravicius, Rafael E. Banchs, Elena Volodina and Courtney Napoles .............. 53

Topicality-Based Indices for Essay Scoring
Beata Beigman Klebanov, Michael Flor and Binod Gyawali ................................ 63

Predicting the Spelling Difficulty of Words for Language Learners
Lisa Beinborn, Torsten Zesch and Iryna Gurevych ........................................ 73

Characterizing Text Difficulty with Word Frequencies
Xiaobin Chen and Detmar Meurers ...................................................... 84

Unsupervised Modeling of Topical Relevance in L2 Learner Text
Ronan Cummins, Helen Yannakoudakis and Ted Briscoe .................................. 95

UW-Stanford System Description for AESW 2016 Shared Task on Grammatical Error Detection
Dan Flickinger, Michael Goodman and Woodley Packard .................................. 105

Shallow Semantic Reasoning from an Incomplete Gold Standard for Learner Language
Levi King and Markus Dickinson .................................................................. 112

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 ........................ 122

Automated scoring across different modalities
Anastassia Loukina and Aoife Cahill ................................................... 130

Model Combination for Correcting Preposition Selection Errors
Nitin Madnani, Michael Heilman and Aoife Cahill ........................................... 136
Pictogrammar: an AAC device based on a semantic grammar
Fernando Martínez-Santiago, Miguel Ángel García Cumbreras, Arturo Montejo Ráez and Manuel Carlos Díaz Galiano ................................................................. 142

Detecting Context Dependence in Exercise Item Candidates Selected from Corpora
Ildikó Pilán ................................................................. 151

Feature-Rich Error Detection in Scientific Writing Using Logistic Regression
Madeline Remse, Mohsen Mesgar and Michael Strube ........................................ 162

Bundled Gap Filling: A New Paradigm for Unambiguous Cloze Exercises
Michael Wojatzki, Oren Melamud and Torsten Zesch ........................................ 172

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 ................................................................. 182

Linguistically Aware Information Retrieval: Providing Input Enrichment for Second Language Learners
Maria Chinkina and Detmar Meurers ................................................................. 188

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 Mulholand ................................................................. 199

Cost-Effectiveness in Building a Low-Resource Morphological Analyzer for Learner Language
Scott Ledbetter and Markus Dickinson ................................................................. 206

Automatically Scoring Tests of Proficiency in Music Instruction
Nitin Madnani, Aoife Cahill and Brian Riordan ................................................................. 217

Combined Tree Kernel-based classifiers for Assessing Quality of Scientific Text
Liliana Mamani Sanchez and Hector-Hugo Franco-Penya ........................................ 223

Augmenting Course Material with Open Access Textbooks
Smitha Milli and Marti A. Hearst ................................................................. 229

Exploring the Intersection of Short Answer Assessment, Authorship Attribution, and Plagiarism Detection
Björn Rudzewitz ................................................................. 235

Sentence-Level Grammatical Error Identification as Sequence-to-Sequence Correction
Allen Schmaltz, Yoon Kim, Alexander M. Rush and Stuart Shieber ........................................ 242

Combining Off-the-shelf Grammar and Spelling Tools for the Automatic Evaluation of Scientific Writing (AESW) Shared Task 2016
René Witte and Bahar Sateli ................................................................. 252
Candidate re-ranking for SMT-based grammatical error correction
Zheng Yuan, Ted Briscoe and Mariano Felice ....................................................... 256

Spoken Text Difficulty Estimation Using Linguistic Features
Su-Youn Yoon, Yeonsuk Cho and Diane Napolitano ........................................... 267

Automatically Extracting Topical Components for a Response-to-Text Writing Assessment
Zahra Rahimi and Diane Litman ........................................................................ 277

Sentence Similarity Measures for Fine-Grained Estimation of Topical Relevance in Learner Essays
Marek Rei and Ronan Cummins ........................................................................ 283

Insights from Russian second language readability classification: complexity-dependent training requirements, and feature evaluation of multiple categories
Robert Reynolds ................................................................................................. 289

Investigating Active Learning for Short-Answer Scoring
Andrea Horbach and Alexis Palmer ...................................................................... 301