

EMNLP 2017

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

Proceedings of the Workshop

September 8, 2017
Copenhagen, Denmark

Gold Sponsors

turnitin | LightSide

 grammarly



Silver Sponsors



Bronze Sponsors



©2017 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 978-1-945626-85-2

Introduction

This has been a momentous year for the BEA Workshop. In its 12th year, the BEA workshop is, for the first time, being held in conjunction with EMNLP. In addition, the workshop is being sponsored by the newly formed Special Interest Group: SIG EDU.¹

Since the first workshop in 1997, BEA has become the leading venue for sharing and 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. The research presented at the workshop highlights advances in the technology and the maturity of the field of NLP in education. The capabilities serve as a response to educational challenges and are poised to support the needs of a variety of stakeholders, including educators, learners, parents, and administrators.

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 and explosive growth of mobile applications for game-based applications for instruction and assessment. The current educational and assessment landscape 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.

Steady growth in the development of NLP-based applications for education has prompted an increased number of workshops that typically focus on a single subfield. In BEA, we make an effort to have papers from many subfields, for example, tools for automated scoring, automated test-item generation, curriculum development, evaluation of text, dialogue, evaluation of genres beyond essays, feedback studies, and grammatical error correction.

This year we received a record 62 submissions, and accepted 9 papers as oral presentations and 25 as poster presentation and/or demos, for an overall acceptance rate of 55 percent. 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.

The BEA12 workshop has presentations on Automated Writing Evaluation (AWE), item generation,

¹https://www.aclweb.org/adminwiki/index.php?title=2017Q3_Reports:_SIGEDU

readability, dialogue and annotation/database schemas, among others:

AWE Written Assessments: Whereas much work in scoring at BEA focuses on learner language, Horbach et al. score essays written by proficient native German speakers in a complex writing task. Madnani et al. look at scoring for content in science, math, language arts and social studies. Rei looks at detecting off-topic essay responses to visual prompts. Riordan et al. examine neural architectures for scoring responses to short answer questions. Finally, looking at the bigger picture, Burstein et al. explore the relations between AWE and broader educational outcomes.

Domain-Specific AWE: Three papers look at assessments in specific subject domains. For language learning, Tolmachev and Kurohashi extract exemplar sentences to accompany flash cards. Tack et al. investigate the feasibility of automated learner English assessment in the CEFR (European) framework. In the science domain, Nadeem and Ostendorf look at language-based mapping of science assessment items to skills.

Error Detection and Correction: Rei and Yannakoudakis use a neural sequence labeling approach to grammatical error detection. Napoles and Callison-Burch adapt Machine Translation (MT) to grammatical error correction. In another use for machine translation, Rei et al. use MT to generate artificial errors for training machine learning systems. Chollampatt and Ng augment an MT approach with neural network models. Farag et al. develop an error-oriented word embedding approach that exploits errors in learner productions. Caines et al. collect crowd-sourced fluency corrections for transcripts of spoken learner English. Finally, Sakaguchi et al. present a position paper on error correction that discusses issues that need to be addressed and provide recommendations.

Item generation: Jiang and Lee develop distractors for fill-in-the-gap items in Chinese. Satria and Tokunaga evaluate automatically generated pronoun reference questions. Chinkina and Meurers generate questions for evaluating language learning. Finally, Stasaski and Hearst generate multiple choice questions using an ontology.

Estimating Item Difficulty: A last topic in the test domain is Pado's paper on estimating question difficulty in the domain of automatic grading.

Readability: Gonzalez-Garduño and Søgaaard measure gaze to predict readability while Štajner et al. measure viewing time per word in autistic and neurotypical readers. Yaneva et al. also explore readability assessment for people with cognitive disabilities. Beigman Klebanov et al. study the challenges of varying text complexity in a read-aloud intervention program. Östling and Grigonyte use deep convolutional neural networks to measure text quality. Sheng et al. introduce the pedagogical roles of documents to study pedagogical values. Gordon et al. generate reading lists of technical text. Finally, Wolska and Clausen simplify metaphorical language for young readers.

Dialogue: There are two papers on dialogue, but with very different topics. In the first, Lugini and Litman predict specificity in classroom discussions. In the second, Jin et al. develop a system for interpreting questions in a virtual patient dialogue system.

Annotation/Databases: Loughnane et al. create a database that links learning content, linguistic annotation and open-source resources. Laarmann-Quante et al. develop a novel German learner corpus.

Finally there are two papers with content so original that they don't fit into any of the categories above: Kochmar and Shutova investigate how semantic knowledge is acquired in English as a second language and evaluate the pace of development across a number of dimensions. Chen and Lee predict an audience's

laughter during an oral presentation.

This year, the workshop is hosting a Shared Task on Native Language Identification² (NLI). NLI is the process of automatically identifying the native language (L1) of a non-native speaker based solely on language that he or she produces in another language. Two previous shared tasks on NLI have been organized in which the task was to identify the native language of non-native speakers of English based on essays and spoken responses to a standardized assessment of academic English proficiency. The first shared task³ was based on the essays only and was also held with the BEA workshop in 2013. Three years later, Computational Paralinguistics Challenge⁴ at Interspeech 2016 hosted a sub-challenge on identifying the native language based solely on the *spoken* responses. This year's shared task combines the inputs from the two previous tasks. There are three tracks: NLI on the essay only, NLI on the speech response only, and NLI using both responses from a test taker. 19 teams competed in the NLI shared task, with 17 presenting their systems during the poster session. A summary report of the shared task (Malmasi et al.) 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 is attending this workshop. We would especially like to thank our sponsors: at the Gold Level, Turnitin | LightSide, Grammarly and Duolingo; at the Silver level, Educational Testing Service (ETS), Pacific Metrics, National Board of Medical Examiners (NBME), and iLexIR; at the Bronze level, Cognii. Their contributions help fund workshop extras, such as the dinner which is a great social and networking event, especially for students.

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

²<https://sites.google.com/site/nlishopredtask/home>

³<https://sites.google.com/site/nlishopredtask2013/home>

⁴<http://emotion-research.net/sigs/speech-sig/is16-compare>

Organizers:

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

Program Committee:

David Alfter, University of Gothenburg
Dimitrios Alikaniotis, University of Cambridge
Øistein E. Andersen, University of Cambridge
Rafael E. Banchs, Institute for Infocomm Research
Rajendra Banjade, University of Memphis
Timo Baumann, Carnegie Mellon University
Lee Becker, Hapara
Beata Beigman Klebanov, Educational Testing Service
Lisa Beinborn, Technische Universität Darmstadt
Kay Berkling, Cooperative State University, Karlsruhe
Delphine Bernhard, LiLPa, Université de Strasbourg
Yevgeni Berzak, MIT
Sameer Bhatnagar, Polytechnique Montreal
Serge Bibauw, KU Leuven & Université Catholique de Louvain
Joachim Bingel, University of Copenhagen
Kristy Boyer, University of Florida
Chris Brew, Digital Operatives LLC
Ted Briscoe, University of Cambridge
Chris Brockett, Microsoft Research
Julian Brooke, University of Melbourne
Dominique Brunato, Istituto di Linguistica Computazionale (ILC-CNR), National Council of Research, Pisa
Christopher Bryant, University of Cambridge
Aoife Cahill, Educational Testing Service
Andrew Caines, University of Cambridge
Lei Chen, Educational Testing Service
Wei-Fan Chen, Academia Sinica
Xioabin Chen, Tübingen University
Yeonsuk Cho, Educational Testing Service
Martin Chodorow, City University of New York
Shamil Chollampatt, National University of Singapore
Mark Core, University of Southern California
Scott Crossley, Georgia State University
Ronan Cummins, University of Cambridge
Luis Fernando D'Haro, Institute for Infocomm Research, A*STAR
Vidas Daudaravicius, VTEX Research
Markus Dickinson, Indiana University
Yo Ehara, National Institute of Advanced Industrial Science and Technology
Keelan Evanini, Educational Testing Service

Mariano Felice, University of Cambridge
Michael Flor, Educational Testing Service
Peter Foltz, Pearson
Hector-Hugo Franco-Penya, DIT
Thomas François, Université Catholique de Louvain
Michael Gamon, Microsoft Research
Dipesh Gautam, University of Memphis
Binyam Gebrekidan Gebre, Philips
Kallirroi Georgila, University of Southern California
Cyril Goutte, National Research Council Canada
Roman Grundkiewicz, Adam Mickiewicz University in Poznan
Iryna Gurevych, TU Darmstadt
Na-Rae Han, University of Pittsburgh
Jiangang Hao, Educational Testing Service
Rachel Harsley, University of Illinois at Chicago
Homa B. Hashemi, University of Pittsburgh
Marti A. Hearst, UC Berkeley
Trude Heift, Simon Fraser University
Derrick Higgins, American Family Insurance
Andrea Horbach, University of Duisburg-Essen
Chung-Chi Huang, Frostburg State University
Radu Tudor Ionescu, University of Bucharest
Ross Israel, Factual
Pamela Jordan, University of Pittsburgh
Marcin Junczys-Dowmunt, Adam Mickiewicz University in Poznan
Fazel Keshtkar, St. John's University
Levi King, Indiana University
Sigrid Klerke, University of Copenhagen
Ekaterina Kochmar, University of Cambridge
Mamoru Komachi, Tokyo Metropolitan University
Robert Krovetz, Lexical Research
Girish Kumar, Stanford University
Lun-Wei Ku, IIS, Academia Sinica
Kristopher Kyle, University of Hawaii
John Lee, City University of Hong Kong
Lung-Hao Lee, National Taiwan Normal University
James Lester, North Carolina State University
Chen Wee Leong, Educational Testing Service
Baoli Li, Henan University of Technology
Diane Litman, University of Pittsburgh
Annie Louis, University of Essex
Anastassia Loukina, Educational Testing Service
Xiaofei Lu, Pennsylvania State University
Fabiana MacMillan, Rosetta Stone
Nitin Madnani, Educational Testing Service
Shervin Malmasi, Harvard Medical School
Liliana Mamani Sanchez, University College Dublin
Montse Maritxalar, University of the Basque Country
Ditty Matthew, Indian Institute of Technology Madras
Julie Medero, Harvey Mudd College
Beata Megyesi, Uppsala University

Mohsen Mesgar, Heidelberg Institute for Theoretical Studies
Angeliki Metallinou, Amazon
Detmar Meurers, University of Tübingen
Christian M. Meyer, Technische Universität Darmstadt
Lisa Michaud, Aspect Software
Michael Mohler, Language Computer Corporation
Maria Moritz, Faculty of Mathematics and Computer Science
Smaranda Muresan, Columbia University
William Murray, Pearson
Courtney Napoles, Johns Hopkins University
Hwee Tou Ng, National University of Singapore
Huy Nguyen, University of Pittsburgh
Nobal Niraula, Boeing Research & Technology
Simon Ostermann, Saarland University
Alexis Palmer, University of North Texas
Siddharth Patwardhan, Apple
Ted Pedersen, University of Minnesota, Duluth
Isaac Persing, University of Texas at Dallas
Patti Price, PPRICE Speech and Language Technology Consulting
Martí Quixal, Universitat Oberta de Catalunya
Preethi Raghavan, IBM
Zahra Rahimi, University of Pittsburgh
Lakshmi Ramachandran, A9.com
David Randolph, University of Illinois at Chicago
Sudha Rao, University of Maryland
Livy Real, IBM Research
Marek Rei, University of Cambridge
Robert Reynolds, UiT The Arctic University of Norway
Brian Riordan, Educational Testing Service
Andrew Rosenberg, IBM
Alla Rozovskaya, Queens College (CUNY)
Alexander Rush, Harvard
Anton Rytting, University of Maryland
Keisuke Sakaguchi, Johns Hopkins University
Elizabeth Salesky, MIT LL
Allen Schmaltz, Harvard University
Swapna Somasundaran, Educational Testing Service
Helmer Strik, Radboud University Nijmegen
David Suendermann-Oeft, Educational Testing Service
Kaveh Taghipour, National University of Singapore
Fatemeh Torabi Asr, Indiana University
Yuen-Hsien Tseng, National Taiwan Normal University
Sowmya Vajjala, Iowa State University
Giulia Venturi, Istituto di Linguistica Computazionale (ILC-CNR)
Aline Villavicencio, Federal University of Rio Grande do Sul
Carl Vogel, Computational Linguistics Lab, O'Reilly Institute
Elena Volodina, University of Gothenburg
Michael White, The Ohio State University
Denise Whitelock, The Open University
David Wible, National Central University
Alistair Willis, The Open University

Magdalena Wolska, Eberhard Karls Universität Tübingen
Michael Wojatzki, University of Duisburg-Essen
Huichao Xue, Google
Zheng Yuan, University of Cambridge
Marcos Zampieri, Saarland University
Torsten Zesch, University of Duisburg-Essen
Fan Zhang, University of Pittsburgh
Xiaodan Zhu, National Research Council Canada

Table of Contents

<i>Question Difficulty – How to Estimate Without Norming, How to Use for Automated Grading</i> Ulrike Pado	1
<i>Combining CNNs and Pattern Matching for Question Interpretation in a Virtual Patient Dialogue System</i> Lifeng Jin, Michael White, Evan Jaffe, Laura Zimmerman and Douglas Danforth	11
<i>Continuous fluency tracking and the challenges of varying text complexity</i> Beata Beigman Klebanov, Anastassia Loukina, John Sabatini and Tenaha O’Reilly	22
<i>Auxiliary Objectives for Neural Error Detection Models</i> Marek Rei and Helen Yannakoudakis	33
<i>Linked Data for Language-Learning Applications</i> Robyn Loughnane, Kate McCurdy, Peter Kolb and Stefan Selent	44
<i>Predicting Specificity in Classroom Discussion</i> Luca Lugini and Diane Litman	52
<i>A Report on the 2017 Native Language Identification Shared Task</i> Shervin Malmasi, Keelan Evanini, Aoife Cahill, Joel Tetreault, Robert Pugh, Christopher Hamill, Diane Napolitano and Yao Qian	62
<i>Evaluation of Automatically Generated Pronoun Reference Questions</i> Arief Yudha Satria and Takenobu Tokunaga	76
<i>Predicting Audience’s Laughter During Presentations Using Convolutional Neural Network</i> Lei Chen and Chong Min Lee	86
<i>Collecting fluency corrections for spoken learner English</i> Andrew Caines, Emma Flint and Paula Buttery	91
<i>Exploring Relationships Between Writing & Broader Outcomes With Automated Writing Evaluation</i> Jill Burstein, Dan McCaffrey, Beata Beigman Klebanov and Guangming Ling	101
<i>An Investigation into the Pedagogical Features of Documents</i> Emily Sheng, Prem Natarajan, Jonathan Gordon and Gully Burns	109
<i>Combining Multiple Corpora for Readability Assessment for People with Cognitive Disabilities</i> Victoria Yaneva, Constantin Orasan, Richard Evans and Omid Rohanian	121
<i>Automatic Extraction of High-Quality Example Sentences for Word Learning Using a Determinantal Point Process</i> Arseny Tolmachev and Sadao Kurohashi	133
<i>Distractor Generation for Chinese Fill-in-the-blank Items</i> Shu Jiang and John Lee	143
<i>An Error-Oriented Approach to Word Embedding Pre-Training</i> Youmna Farag, Marek Rei and Ted Briscoe	149
<i>Investigating neural architectures for short answer scoring</i> Brian Riordan, Andrea Horbach, Aoife Cahill, Torsten Zesch and Chong Min Lee	159

<i>Human and Automated CEFR-based Grading of Short Answers</i> Anaïs Tack, Thomas François, Sophie Roekhaut and Cédric Fairon	169
<i>GEC into the future: Where are we going and how do we get there?</i> Keisuke Sakaguchi, Courtney Napoles and Joel Tetreault	180
<i>Detecting Off-topic Responses to Visual Prompts</i> Marek Rei	188
<i>Combining Textual and Speech Features in the NLI Task Using State-of-the-Art Machine Learning Techniques</i> Pavel Ircing, Jan Svec, Zbynek Zajic, Barbora Hladka and Martin Holub	198
<i>Native Language Identification Using a Mixture of Character and Word N-grams</i> Elham Mohammadi, Hadi Veisi and Hessam Amini	210
<i>Ensemble Methods for Native Language Identification</i> Sophia Chan, Maryam Honari Jahromi, Benjamin Benetti, Aazim Lakhani and Alona Fyshe ..	217
<i>Can string kernels pass the test of time in Native Language Identification?</i> Radu Tudor Ionescu and Marius Popescu	224
<i>Neural Networks and Spelling Features for Native Language Identification</i> Johannes Bjerva, Gintare Grigonyte, Robert Östling and Barbara Plank	235
<i>A study of N-gram and Embedding Representations for Native Language Identification</i> Sowmya Vajjala and Sagnik Banerjee	240
<i>A Shallow Neural Network for Native Language Identification with Character N-grams</i> Yunita Sari, Muhammad Rifqi Fatchurrahman and Meisyarah Dwiastuti	249
<i>Fewer features perform well at Native Language Identification task</i> Taraka Rama and Çağrı Çöltekin	255
<i>Structured Generation of Technical Reading Lists</i> Jonathan Gordon, Stephen Aguilar, Emily Sheng and Gully Burns	261
<i>Effects of Lexical Properties on Viewing Time per Word in Autistic and Neurotypical Readers</i> Sanja Štajner, Victoria Yaneva, Ruslan Mitkov and Simone Paolo Ponzetto	271
<i>Transparent text quality assessment with convolutional neural networks</i> Robert Östling and Gintare Grigonyte	282
<i>Artificial Error Generation with Machine Translation and Syntactic Patterns</i> Marek Rei, Mariano Felice, Zheng Yuan and Ted Briscoe	287
<i>Modelling semantic acquisition in second language learning</i> Ekaterina Kochmar and Ekaterina Shutova	293
<i>Multiple Choice Question Generation Utilizing An Ontology</i> Katherine Stasaski and Marti A. Hearst	303
<i>Simplifying metaphorical language for young readers: A corpus study on news text</i> Magdalena Wolska and Yulia Clausen	313

<i>Language Based Mapping of Science Assessment Items to Skills</i> Farah Nadeem and Mari Ostendorf	319
<i>Connecting the Dots: Towards Human-Level Grammatical Error Correction</i> Shamil Chollampatt and Hwee Tou Ng	327
<i>Question Generation for Language Learning: From ensuring texts are read to supporting learning</i> Maria Chinkina and Detmar Meurers	334
<i>Systematically Adapting Machine Translation for Grammatical Error Correction</i> Courtney Napoles and Chris Callison-Burch	345
<i>Fine-grained essay scoring of a complex writing task for native speakers</i> Andrea Horbach, Dirk Scholten-Akoun, Yuning Ding and Torsten Zesch	357
<i>Exploring Optimal Voting in Native Language Identification</i> Cyril Goutte and Serge Léger	367
<i>CIC-FBK Approach to Native Language Identification</i> Iliia Markov, Lingzhen Chen, Carlo Strapparava and Grigori Sidorov	374
<i>The Power of Character N-grams in Native Language Identification</i> Artur Kulmizev, Bo Blankers, Johannes Bjerva, Malvina Nissim, Gertjan van Noord, Barbara Plank and Martijn Wieling	382
<i>Classifier Stacking for Native Language Identification</i> Wen Li and Liang Zou	390
<i>Native Language Identification on Text and Speech</i> Marcos Zampieri, Alina Maria Ciobanu and Liviu P. Dinu	398
<i>Native Language Identification using Phonetic Algorithms</i> Charese Smiley and Sandra Kübler	405
<i>A deep-learning based native-language classification by using a latent semantic analysis for the NLI Shared Task 2017</i> Yoo Rhee Oh, Hyung-Bae Jeon, Hwa Jeon Song, Yun-Kyung Lee, Jeon-Gue Park and Yun-Keun Lee	413
<i>Fusion of Simple Models for Native Language Identification</i> Fabio Kepler, Ramón Astudillo and Alberto Abad	423
<i>Stacked Sentence-Document Classifier Approach for Improving Native Language Identification</i> Andrea Cimino and Felice Dell’Orletta	430
<i>Using Gaze to Predict Text Readability</i> Ana Valeria Gonzalez-Garduño and Anders Søgaard	438
<i>Annotating Orthographic Target Hypotheses in a German L1 Learner Corpus</i> Ronja Laarmann-Quante, Katrin Ortmann, Anna Ehlert, Maurice Vogel and Stefanie Dipper ..	444
<i>A Large Scale Quantitative Exploration of Modeling Strategies for Content Scoring</i> Nitin Madhani, Anastassia Loukina and Aoife Cahill	457

Workshop Program

Friday, September 8, 2017

8:45–9:00 *Load Oral Presentations*

9:00–10:30 **Session 1**

9:00–9:15 *Opening Remarks*

9:15–9:40 *Question Difficulty – How to Estimate Without Norming, How to Use for Automated Grading*
Ulrike Pado

9:40–10:05 *Combining CNNs and Pattern Matching for Question Interpretation in a Virtual Patient Dialogue System*
Lifeng Jin, Michael White, Evan Jaffe, Laura Zimmerman and Douglas Danforth

10:05–10:30 *Continuous fluency tracking and the challenges of varying text complexity*
Beata Beigman Klebanov, Anastassia Loukina, John Sabatini and Tenaha O'Reilly

10:30–11:00 **Break**

11:00–12:35 **Session 2**

11:00–11:25 *Auxiliary Objectives for Neural Error Detection Models*
Marek Rei and Helen Yannakoudakis

11:25–11:50 *Linked Data for Language-Learning Applications*
Robyn Loughnane, Kate McCurdy, Peter Kolb and Stefan Selent

11:50–12:10 *Predicting Specificity in Classroom Discussion*
Luca Lugini and Diane Litman

12:10–12:35 *A Report on the 2017 Native Language Identification Shared Task*
Shervin Malmasi, Keelan Evanini, Aoife Cahill, Joel Tetreault, Robert Pugh, Christopher Hamill, Diane Napolitano and Yao Qian

12:35–14:00 **Lunch**

Friday, September 8, 2017 (continued)

14:00–15:30 Poster Session

14:00–14:45 *Poster Session A*

Evaluation of Automatically Generated Pronoun Reference Questions

Arief Yudha Satria and Takenobu Tokunaga

Predicting Audience's Laughter During Presentations Using Convolutional Neural Network

Lei Chen and Chong Min Lee

Collecting fluency corrections for spoken learner English

Andrew Caines, Emma Flint and Paula Buttery

Exploring Relationships Between Writing & Broader Outcomes With Automated Writing Evaluation

Jill Burstein, Dan McCaffrey, Beata Beigman Klebanov and Guangming Ling

An Investigation into the Pedagogical Features of Documents

Emily Sheng, Prem Natarajan, Jonathan Gordon and Gully Burns

Combining Multiple Corpora for Readability Assessment for People with Cognitive Disabilities

Victoria Yaneva, Constantin Orasan, Richard Evans and Omid Rohanian

Automatic Extraction of High-Quality Example Sentences for Word Learning Using a Determinantal Point Process

Arseny Tolmachev and Sadao Kurohashi

Distractor Generation for Chinese Fill-in-the-blank Items

Shu Jiang and John Lee

An Error-Oriented Approach to Word Embedding Pre-Training

Youmna Farag, Marek Rei and Ted Briscoe

Investigating neural architectures for short answer scoring

Brian Riordan, Andrea Horbach, Aoife Cahill, Torsten Zesch and Chong Min Lee

Human and Automated CEFR-based Grading of Short Answers

Anaïs Tack, Thomas François, Sophie Roekhaut and Cédric Fairon

GEC into the future: Where are we going and how do we get there?

Keisuke Sakaguchi, Courtney Napoles and Joel Tetreault

Detecting Off-topic Responses to Visual Prompts

Marek Rei

Friday, September 8, 2017 (continued)

Combining Textual and Speech Features in the NLI Task Using State-of-the-Art Machine Learning Techniques

Pavel Ircing, Jan Svec, Zbynek Zajic, Barbora Hladka and Martin Holub

Native Language Identification Using a Mixture of Character and Word N-grams

Elham Mohammadi, Hadi Veisi and Hessam Amini

Ensemble Methods for Native Language Identification

Sophia Chan, Maryam Honari Jahromi, Benjamin Benetti, Aazim Lakhani and Alona Fyshe

Can string kernels pass the test of time in Native Language Identification?

Radu Tudor Ionescu and Marius Popescu

Neural Networks and Spelling Features for Native Language Identification

Johannes Bjerva, Gintare Grigonyte, Robert Östling and Barbara Plank

A study of N-gram and Embedding Representations for Native Language Identification

Sowmya Vajjala and Sagnik Banerjee

A Shallow Neural Network for Native Language Identification with Character N-grams

Yunita Sari, Muhammad Rifqi Fatchurrahman and Meisyarah Dwiastuti

Fewer features perform well at Native Language Identification task

Taraka Rama and Çağrı Çöltekin

14:45–15:30 *Poster Session B*

Structured Generation of Technical Reading Lists

Jonathan Gordon, Stephen Aguilar, Emily Sheng and Gully Burns

Effects of Lexical Properties on Viewing Time per Word in Autistic and Neurotypical Readers

Sanja Štajner, Victoria Yaneva, Ruslan Mitkov and Simone Paolo Ponzetto

Transparent text quality assessment with convolutional neural networks

Robert Östling and Gintare Grigonyte

Artificial Error Generation with Machine Translation and Syntactic Patterns

Marek Rei, Mariano Felice, Zheng Yuan and Ted Briscoe

Modelling semantic acquisition in second language learning

Ekaterina Kochmar and Ekaterina Shutova

Multiple Choice Question Generation Utilizing An Ontology

Katherine Stasaski and Marti A. Hearst

Friday, September 8, 2017 (continued)

Simplifying metaphorical language for young readers: A corpus study on news text
Magdalena Wolska and Yulia Clausen

Language Based Mapping of Science Assessment Items to Skills
Farah Nadeem and Mari Ostendorf

Connecting the Dots: Towards Human-Level Grammatical Error Correction
Shamil Chollampatt and Hwee Tou Ng

Question Generation for Language Learning: From ensuring texts are read to supporting learning
Maria Chinkina and Detmar Meurers

Systematically Adapting Machine Translation for Grammatical Error Correction
Courtney Napoles and Chris Callison-Burch

Fine-grained essay scoring of a complex writing task for native speakers
Andrea Horbach, Dirk Scholten-Akoun, Yuning Ding and Torsten Zesch

Exploring Optimal Voting in Native Language Identification
Cyril Goutte and Serge Léger

CIC-FBK Approach to Native Language Identification
Iliia Markov, Lingzhen Chen, Carlo Strapparava and Grigori Sidorov

The Power of Character N-grams in Native Language Identification
Artur Kulmizev, Bo Blankers, Johannes Bjerva, Malvina Nissim, Gertjan van Noord, Barbara Plank and Martijn Wieling

Classifier Stacking for Native Language Identification
Wen Li and Liang Zou

Native Language Identification on Text and Speech
Marcos Zampieri, Alina Maria Ciobanu and Liviu P. Dinu

Native Language Identification using Phonetic Algorithms
Charese Smiley and Sandra Kübler

A deep-learning based native-language classification by using a latent semantic analysis for the NLI Shared Task 2017
Yoo Rhee Oh, Hyung-Bae Jeon, Hwa Jeon Song, Yun-Kyung Lee, Jeon-Gue Park and Yun-Keun Lee

Fusion of Simple Models for Native Language Identification
Fabio Kepler, Ramón Astudillo and Alberto Abad

Stacked Sentence-Document Classifier Approach for Improving Native Language Identification
Andrea Cimino and Felice Dell'Orletta

Friday, September 8, 2017 (continued)

15:30–16:00 Break

16:00–17:30 Session 3

16:00–16:25 *Using Gaze to Predict Text Readability*
Ana Valeria Gonzalez-Garduño and Anders Søgaard

16:25–16:50 *Annotating Orthographic Target Hypotheses in a German LI Learner Corpus*
Ronja Laarmann-Quante, Katrin Ortmann, Anna Ehlert, Maurice Vogel and Stefanie Dipper

16:50–17:15 *A Large Scale Quantitative Exploration of Modeling Strategies for Content Scoring*
Nitin Madnani, Anastassia Loukina and Aoife Cahill

17:15–17:30 Closing Remarks

