

NAACL HLT 2010

**Fifth Workshop on
Innovative Use of NLP for
Building Educational
Applications**

Proceedings of the Workshop

June 5, 2010
Los Angeles, California

USB memory sticks produced by
Omnipress Inc.
2600 Anderson Street
Madison, WI 53707
USA

©2010 The Association for Computational Linguistics

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

Introduction

NLP researchers are now building educational applications across a number of areas, including automated evaluation of student writing and speaking, rich grammatical error detection with an increasing focus on English language learning, tools to support student reading, and intelligent tutoring.

This workshop is the fifth in a series, specifically related to Building NLP Applications for Education, that began at NAACL/HLT (2003), and continued at ACL 2005 (Ann Arbor), ACL/HLT 2008 (Columbus), NAACL/HLT 2009 (Boulder), and now at NAACL/HLT 2010 (Los Angeles). Research in this area continues to grow, and there is ever-increasing interest and practical application which was evidenced this year, again, by an even larger number of submissions.

We received a record 28 submissions and accepted 13 papers, two of which include demos. All of the papers are published in these proceedings. Each paper was carefully reviewed by two members of the Program Committee. We selected reviewers most appropriate for each paper so as to give more helpful feedback and comments. This workshop offers an opportunity to present and publish work that is highly relevant to NAACL, but is also highly specialized, and so this workshop is often a more appropriate venue for such work. While the field is growing, we do recognize that there is a core group of institutions and researchers who work in this area. That said, we continue to have a very strong policy to deal with conflicts of interest. First, reviewers were not assigned any papers to evaluate if the paper had an author from their institution. Second, with respect to the organizing committee, authors of papers where there was a conflict of interest recused themselves from the discussion.

The papers accepted to this workshop were selected on the basis of several factors: the strength of the research, the novelty of the approach or domain, and the appropriateness for this workshop. The final set of papers fall under several main themes which we show below in the order of the workshop program.

Technology designed to support reading comprehension:

- Readability Assessment for Text Simplification (Aluisio, Specia, Gasperin and Scarton)
- Enhancing Authentic Web Pages for Language Learners (Meurers, Ziai, Amaral, Boyd, Dimitrov, Metcalf and Ott)
- AutoTutor: a piece of cake for teachers (Quixal, Preu, Garca-Narbona and Boullosa)

Learner error detection and annotation:

- Annotating ESL Errors: Challenges and Rewards (Rozovskaya and Roth)
- Search right and thou shalt find ... Using Web Queries for Learner Error Detection (Gamon and Leacock)
- Rethinking Grammatical Error Annotation and Evaluation with the Amazon Mechanical Turk (Tetreault, Filatova and Chodorow)

Classroom assessment and instruction:

- Predicting Cloze Task Quality for Vocabulary Training (Skory and Eskenazi)
- Generating Quantifiers and Negation to Explain Homework Testing (Perry and Shan)
- Leveraging Hidden Dialogue State to Select Tutorial Moves (Boyer, Phillips, Ha, Wallis, Vouk and Lester)

Evaluation of Noisy Data:

- Towards Using Structural Events To Assess Non-native Speech (Lei Chen, Joel Tetreault and Xiaoming Xi)
- A Human-Computer Collaboration Approach to Improve Accuracy of an Automated English Scoring System (Jee Eun Kim and Kong Joo Lee)
- Towards Identifying Unresolved Discussions in Student Online Forums (Jihie Kim, Jia Li and Taehwan Kim)
- Off-topic essay detection using short prompt texts (Louis and Higgins)

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. All of these factors contribute to a truly rich and successful event!

Joel Tetreault, Educational Testing Service
Jill Burstein, Educational Testing Service
Claudia Leacock, Butler-Hill Group

Organizers:

Joel Tetreault, Educational Testing Service
Jill Burstein, Educational Testing Service
Claudia Leacock, Butler Hill Group

Program Committee:

Delphine Bernhard, LIMSI-CNRS, France
Jared Bernstein, Pearson, USA
Martin Chodorow, Hunter College, CUNY, USA
Barbara Di Eugenio, University of Illinois at Chicago, USA
Markus Dickinson, Indiana University, USA
William B. Dolan, Microsoft, USA
Maxine Eskenazi, Carnegie Mellon University, USA
Peter Foltz, Pearson, USA
Jennifer Foster, Dublin City University, Ireland
Annette Frank, University of Heidelberg, Germany
Michael Gamon, Microsoft, USA
Caroline Gasperin, University of Sao Paulo, Brazil
Iryna Gurevych, University of Darmstadt, Germany
Na-Rae Han, University of Pittsburgh, USA
Trude Heift, Simon Frasier University, Canada
Derrick Higgins, ETS, USA
Emi Izumi, NICT, Japan
Pamela Jordan, University of Pittsburgh, USA
Ola Knutsson, KTH Nada, Sweden
John Lee, City University of Hong Kong, China
Diane Litman, University of Pittsburgh, USA
Detmar Meurers, University of Tübingen, Germany
Lisa Michaud, Saint Anselm College, USA
Ani Nenkova, University of Pennsylvania, USA
Mari Ostendorf, University of Washington, USA
Ted Pedersen, University of Minnesota, USA
Mihai Rotaru, TextKernel, the Netherlands
Dan Roth, UIUC, USA
Mathias Schulze, University of Waterloo, Canada
Stephanie Seneff, MIT, USA
Richard Sproat, Oregon Graduate Institute, USA
Jana Sukkarieh, ETS, USA
Svetlana Stoyanchev, Open University, UK
Nai-Lung Tsao, National Central University, Taiwan

Pete Whitelock, Oxford University Press, UK
David Wible, National Central University, Taiwan
Magdalena Wolska, Saarbruken University, Germany

Table of Contents

| | |
|--|----|
| <i>Readability Assessment for Text Simplification</i> Sandra Aluisio, Lucia Specia, Caroline Gasperin and Carolina Scarton | 1 |
| <i>Enhancing Authentic Web Pages for Language Learners</i> Detmar Meurers, Ramon Ziai, Luiz Amaral, Adriane Boyd, Aleksandar Dimitrov, Vanessa Metcalf and Niels Ott | 10 |
| <i>AutoLearn's authoring tool: a piece of cake for teachers</i> Martí Quixal, Susanne Preuß, David García-Narbona and Beto Boullosa | 19 |
| <i>Annotating ESL Errors: Challenges and Rewards</i> Alla Rozovskaya and Dan Roth | 28 |
| <i>Search right and thou shalt find ... Using Web Queries for Learner Error Detection</i> Michael Gamon and Claudia Leacock | 37 |
| <i>Rethinking Grammatical Error Annotation and Evaluation with the Amazon Mechanical Turk</i> Joel Tetreault, Elena Filatova and Martin Chodorow | 45 |
| <i>Predicting Cloze Task Quality for Vocabulary Training</i> Adam Skory and Maxine Eskenazi | 49 |
| <i>Generating Quantifiers and Negation to Explain Homework Testing</i> Jason Perry and Chung-chieh Shan | 57 |
| <i>Leveraging Hidden Dialogue State to Select Tutorial Moves</i> Kristy Boyer, Rob Phillips, Eun Young Ha, Michael Wallis, Mladen Vouk and James Lester ... | 66 |
| <i>Towards Using Structural Events To Assess Non-native Speech</i> Lei Chen, Joel Tetreault and Xiaoming Xi | 74 |
| <i>A Human-Computer Collaboration Approach to Improve Accuracy of an Automated English Scoring System</i> Jee Eun Kim and Kong Joo Lee | 80 |
| <i>Towards Identifying Unresolved Discussions in Student Online Forums</i> Jihie Kim, Jia Li and Taehwan Kim | 84 |
| <i>Off-topic essay detection using short prompt texts</i> Annie Louis and Derrick Higgins | 92 |

Workshop Program

Saturday, June 5, 2010

- 9:00–9:15 Opening Remarks
- 9:15–9:40 *Readability Assessment for Text Simplification*
Sandra Aluisio, Lucia Specia, Caroline Gasperin and Carolina Scarton
- 9:40–10:05 *Enhancing Authentic Web Pages for Language Learners*
Detmar Meurers, Ramon Ziai, Luiz Amaral, Adriane Boyd, Aleksandar Dimitrov,
Vanessa Metcalf and Niels Ott
- 10:05–10:30 *AutoLearn’s authoring tool: a piece of cake for teachers*
Martí Quixal, Susanne Preuß, David García-Narbona and Beto Boulosa
- 10:30–11:00 **Break**
- 11:00–11:25 *Annotating ESL Errors: Challenges and Rewards*
Alla Rozovskaya and Dan Roth
- 11:25–11:50 *Search right and thou shalt find ... Using Web Queries for Learner Error Detection*
Michael Gamon and Claudia Leacock
- 11:50–12:10 *Rethinking Grammatical Error Annotation and Evaluation with the Amazon Mechanical Turk*
Joel Tetreault, Elena Filatova and Martin Chodorow
- 12:10–1:45 **Lunch**
- 1:45–2:10 *Predicting Cloze Task Quality for Vocabulary Training*
Adam Skory and Maxine Eskenazi
- 2:10–2:35 *Generating Quantifiers and Negation to Explain Homework Testing*
Jason Perry and Chung-chieh Shan
- 2:35–3:00 *Leveraging Hidden Dialogue State to Select Tutorial Moves*
Kristy Boyer, Rob Phillips, Eun Young Ha, Michael Wallis, Mladen Vouk and James Lester
- 3:00–3:30 **Break**

Saturday, June 5, 2010 (continued)

- 3:30–3:55 *Towards Using Structural Events To Assess Non-native Speech*
Lei Chen, Joel Tetreault and Xiaoming Xi
- 3:55–4:15 *A Human-Computer Collaboration Approach to Improve Accuracy of an Automated English Scoring System*
Jee Eun Kim and Kong Joo Lee
- 4:15–4:40 *Towards Identifying Unresolved Discussions in Student Online Forums*
Jihie Kim, Jia Li and Taehwan Kim
- 4:40–5:00 *Off-topic essay detection using short prompt texts*
Annie Louis and Derrick Higgins