8th Workshop on Innovative Use of NLP for Building Educational Applications 2013

Atlanta, Georgia, USA
13 June 2013

ISBN: 978-1-62993-001-5

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© (2013) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2013)

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

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

The Utility of Manual and Automatic Linguistic Error Codes for Identifying Neurodevelopmental Disorders
Eric Morley, Brian Roark and Jan van Santen 1

Shallow Semantic Analysis of Interactive Learner Sentences
Levi King and Markus Dickinson 11

Building a Large Annotated Corpus of Learner English: The NUS Corpus of Learner English
Daniel Dahlmeier, Hwee Tou Ng and Siew Mei Wu 22

Developing and testing a self-assessment and tutoring system
Øistein E. Andersen, Helen Yannakoudakis, Fiona Barker and Tim Parish 32

Automated Essay Scoring for Swedish
Robert Östling, André Smolentzov, Björn Tyrefors Hinnerich and Erik Höglin 42

A Report on the First Native Language Identification Shared Task
Joel Tetreault, Daniel Blanchard and Aoife Cahill 48

Applying Unsupervised Learning To Support Vector Space Model Based Speaking Assessment
Lei Chen 58

Role of Morpho-Syntactic Features in Estonian Proficiency Classification
Sowmya Vajjala and Kaidi Loo 63

Automated Content Scoring of Spoken Responses in an Assessment for Teachers of English
Klaus Zechner and Xinhao Wang 73

Experimental Results on the Native Language Identification Shared Task
Amjad Abu-Jbara, Rahul Jha, Eric Morley and Dragomir Radev 82

VTEX System Description for the NLI 2013 Shared Task
Vidas Daudaravicius 89

Feature Space Selection and Combination for Native Language Identification
Cyril Goutte, Serge Léger and Marine Carpuat 96

Discriminating Non-Native English with 350 Words
John Henderson, Guido Zarrella, Craig Pfeifer and John D. Burger 101

Maximizing Classification Accuracy in Native Language Identification
Scott Jarvis, Yves Bestgen and Steve Pepper 111

Recognizing English Learners' Native Language from Their Writings
Baoli LI 119

NLI Shared Task 2013: MQ Submission
Shervin Malmasi, Sze-Meng Jojo Wong and Mark Dras 124

NAIST at the NLI 2013 Shared Task
Tomoya Mizumoto, Yuta Hayashibe, Keisuke Sakaguchi, Mamoru Komachi and Yuji Matsumoto 134

Cognate and Misspelling Features for Natural Language Identification
Garrett Nicolai, Bradley Hauer, Mohammad Salameh, Lei Yao and Grzegorz Kondrak 140

Exploring Syntactic Representations for Native Language Identification
Ben Swanson 146

Simple Yet Powerful Native Language Identification on TOEFL11
Ching-Yi Wu, Po-Hsiang Lai, Yang Liu and Vincent Ng 152

Prompt-based Content Scoring for Automated Spoken Language Assessment
Keelan Evanini, Shasha Xie and Klaus Zechner 157

Automated Scoring of a Summary-Writing Task Designed to Measure Reading Comprehension
Nitin Madnani, Jill Burstein, John Sabatini and Tenaha O'Reilly 163

Inter-annotator Agreement for Dependency Annotation of Learner Language
Marwa Ragheb and Markus Dickinson 169

Native Language Identification with PPM
Victoria Bobicev 180

Using Other Learner Corpora in the 2013 NLI Shared Task
Julian Brooke and Graeme Hirst 188

Combining Shallow and Linguistically Motivated Features in Native Language Identification
Serhiy Bykh, Sowmya Vajjala, Julia Krivanek and Detmar Meurers 197

Linguistic Profiling based on General–purpose Features and Native Language Identification
Andrea Cimino, Felice Dell'Orletta, Giulia Venturi and Simonetta Montemagni 207

Improving Native Language Identification with TF-IDF Weighting
Binyam Gebrekidan Gebre, Marcos Zampieri, Peter Wittenburg and Tom Heskes 216

Native Language Identification: a Simple n-gram Based Approach
Binod Gyawali, Gabriela Ramirez and Thamar Solorio 224

Feature Engineering in the NLI Shared Task 2013: Charles University Submission Report
Barbora Hladka, Martin Holub and Vincent Kriz 232

Native Language Identification: A Key N-gram Category Approach
Kristopher Kyle, Scott Crossley, Jianmin Dai and Danielle McNamara 242

Using N-gram and Word Network Features for Native Language Identification
Shibamouli Lahiri and Rada Mihalcea 251

LIMSI's participation to the 2013 shared task on Native Language Identification
Thomas Lavergne, Gabriel Illouz, Aurélien Max and Ryo Nagata 260

Native Language Identification using large scale lexical features
André Lynum 266

The Story of the Characters, the DNA and the Native Language
Marius Popescu and Radu Tudor Ionescu 270

Identifying the L1 of non-native writers: the CMU-Haifa system
Yulia Tsvetkov, Naama Twitto, Nathan Schneider, Noam Ordan, Manaal Faruqui, Victor Chahuneau, Shuly Wintner and Chris Dyer 279

Evaluating Unsupervised Language Model Adaptation Methods for Speaking Assessment
Shasha Xie and Lei Chen 288

Improving interpretation robustness in a tutorial dialogue system
Myroslava Dzikovska, Elaine Farrow and Johanna Moore 293

Detecting Missing Hyphens in Learner Text
Aoife Cahill, Martin Chodorow, Susanne Wolff and Nitin Madnani 300

Applying Machine Translation Metrics to Student-Written Translations
Lisa Michaud and Patricia Ann McCoy 306