9th Linguistic Annotation Workshop 2015

(LAW IX)

Held at NAACL HLT 2015

Denver, Colorado, USA
5 June 2015

ISBN: 978-1-5108-0431-9

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© (2015) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2015)

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

Scaling Semantic Frame Annotation
Nancy Chang, Praveen Paritosh, David Huynh and Collin Baker 1

An Analytic and Empirical Evaluation of Return-on-Investment-Based Active Learning
Robbie Haertel, Eric Ringger, Kevin Seppi and Paul Felt 11

Annotating genericity: a survey, a scheme, and a corpus
Annemarie Friedrich, Alexis Palmer, Melissa Peate Sørensen and Manfred Pinkal 21

Design and Annotation of the First Italian Corpus for Text Simplification
Dominique Brunato, Felice Dell'Orletta, Giulia Venturi and Simonetta Montemagni 31

On the Discoursive Structure of Computer Graphics Research Papers
Beatriz Fisas, Horacio Saggion and Francesco Ronzano 42

Semantic Annotation of Japanese Functional Expressions and its Impact on Factuality Analysis
Yudai Kamioka, Kazuya Narita, Junta Mizuno, Miwa Kanno and Kentaro Inui 52

A Qualitative Analysis of a Corpus of Opinion Summaries based on Aspects
Roque Lopez, Thiago Pardo, Lucas Avanço, Pedro Filho, Alessandro Bokan, Paula Cardoso, Márcio Dias, Fernando Nóbrega, Marco Cabezudo, Jackson Souza, Andressa Zacarias, Eloize Seno and Ariani Di Felippo 62

Developing Language-tagged Corpora for Code-switching Tweets
Suraj Maharjan, Elizabeth Blair, Steven Bethard and Thamar Solorio 72

Annotating Geographical Entities on Microblog Text
Koji Matsuda, Akira Sasaki, Naoaki Okazaki and Kentaro Inui 85

The Annotation Process of the ITU Web Treebank
Tuğba Pamay, Umut Sulubacak, Dilara Torunoğlu-Selamet and Gülşen Eryiğit 95

Part of Speech Annotation of Intermediate Versions in the Keystroke Logged Translation Corpus
Tatiana Serbina, Paula Niemietz, Matthias Fricke, Philipp Meisen and Stella Neumann 102

A Hierarchy with, of, and for Preposition Supersenses
Nathan Schneider, Vivek Srikumar, Jena D. Hwang and Martha Palmer 112

Bilingual English-Czech Valency Lexicon Linked to a Parallel Corpus
Zdenka Uresova, Ondřej Dušek, Eva Fucikova, Jan Hajic and Jana Sindlerova 124

Correction Annotation for Non-Native Arabic Texts: Guidelines and Corpus
Wajdi Zaghouani, Nizar Habash, Houda Bouamor, Alla Rozovskaya, Behrang Mohit, Abeer Heider and Kemal Oflazer 129

Balancing the Existing and the New in the Context of Annotating Non-Canonical Language
Ann Bies 140

Parsing Learner Text: to Shoehorn or not to Shoehorn
Aoife Cahill 144

Non-canonical language is not harder to annotate than canonical language
Barbara Plank, Héctor Martínez Alonso and Anders Søgaard 148

What I've learned about annotating informal text (and why you shouldn't take my word for it)
Nathan Schneider 152

On Grammaticality in the Syntactic Annotation of Learner Language
Markus Dickinson and Marwa Ragheb 158

Across Languages and Genres: Creating a Universal Annotation Scheme for Textual Relations
Ekaterina Lapshinova-Koltunski, Anna Nedoluzhko and Kerstin Anna Kunz 168

Annotating the Implicit Content of Sluices
Pranav Anand and Jim McCloskey 178

Annotating Causal Language Using Corpus Lexicography of Constructions
Jesse Dunietz, Lori Levin and Jaime Carbonell 188