8th Linguistic Annotation Workshop 2014

Held at COLING 2014

Dublin, Ireland
23 - 24 August 2014

ISBN: 978-1-63439-219-8

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© (2014) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2014)

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

STTS 2.0? Improving the Tagset for the Part-of-Speech-Tagging of German Spoken Data
Swantje Westpfahl 1

Use of Coreference in Automatic Searching for Multiword Discourse Markers in the Prague Dependency Treebank
Magdalena Rysova and Jiří Mírovský 11

POS Error Detection in Automatically Annotated Corpora
Ines Rehbein 20

Aligning Chinese-English Parallel Parse Trees: Is it Feasible?
Dun Deng and Nianwen Xue 29

Sentence Diagrams: Their Evaluation and Combination
Jirka Hana, Barbora Hladka and Ivana Luksova 38

Finding Your "Inner-Annotator": An Experiment in Annotator Independence for Rating Discourse Coherence Quality in Essays
Jill Burstein, Swapna Somasundaran and Martin Chodorow 48

Optimizing Annotation Efforts to Build Reliable Annotated Corpora for Training Statistical Models
Cyril Grouin, Thomas Lavergne and Aurelie Neveol 54

A Web-based Geo-resolution Annotation and Evaluation Tool
Beatrice Alex, Kate Byrne, Claire Grover and Richard Tobin 59

Annotating Uncertainty in Hungarian Webtext
Veronika Vincze, Katalin Ilona Simkó and Viktor Varga 64

A Corpus Study for Identifying Evidence on Microblogs
Paul Reisert, Junta Mizuno, Miwa Kanno, Naoaki Okazaki and Kentaro Inui 70

Semi-Semantic Part of Speech Annotation and Evaluation
Qaiser Abbas 75

Multiple Views as Aid to Linguistic Annotation Error Analysis
Marilena Di Bari, Serge Sharoff and Martin Thomas 82

Annotating a Large Representative Corpus of Clinical Notes for Parts of Speech
Narayan Choudhary, Parth Pathak, Pinal Patel and Vishal Panchal 87

Part-of-speech Tagset and Corpus Development for Igbo, an African Language
Ikechukwu Onyenwe, Chinedu Uchechukwu and Mark Hepple 93

Annotating Descriptively Incomplete Language phenomena
Fabian Barteld, Sarah Ihden, Ingrid Schröder and Heike Zinsmeister 99

Annotating Discourse Connectives in Spoken Turkish
Isin Demirsahin and Deniz Zeyrek 105

Exploiting the Human Computational Effort Dedicated to Message Reply Formatting for Training Discursive Email Segmenters
Nicolas Hernandez and Soufian Salim 110

Annotating Multiparty Discourse: Challenges for Agreement Metrics
Nina Wacholder, Smaranda Muresan, Debanjan Ghosh and Mark Aakhus 120

Towards Automatic Annotation of Clinical Decision-Making Style
Limor Hochberg, Cecilia Ovesdotter Alm, Esa M. Rantanen, Qi Yu, Caroline M. DeLong and Anne Haake 129

Interactive Annotation for Event Modality in Modern Standard and Egyptian Arabic Tweets
Rania Al-Sabbagh, Roxana Girju and Jana Diesner 139

Situation Entity Annotation
Annemarie Friedrich and Alexis Palmer 149

Focus Annotation in Reading Comprehension Data
Ramon Ziai and Detmar Meurers 159