

Linguistic Annotation Workshop 2007

(The LAW)

Held in conjunction with ACL 2007

**Prague, Czech Republic
28-29 June 2007**

ISBN: 978-1-61738-522-3

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571
www.proceedings.com

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2007) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2010)

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

Table of Contents

<i>GrAF: A Graph-based Format for Linguistic Annotations</i> Nancy Ide and Keith Suderman	1
<i>Efficient Annotation with the Jena ANnotation Environment (JANE)</i> Katrin Tomanek, Joachim Wermter and Udo Hahn	9
<i>Mining Syntactically Annotated Corpora with XQuery</i> Gosse Bouma and Geert Kloosterman	17
<i>Associating Facial Displays with Syntactic Constituents for Generation</i> Mary Ellen Foster	25
<i>An Annotation Type System for a Data-Driven NLP Pipeline</i> Udo Hahn, Ekaterina Buyko, Katrin Tomanek, Scott Piao, John McNaught, Yoshimasa Tsuruoka and Sophia Ananiadou	33
<i>Discontinuity Revisited: An Improved Conversion to Context-Free Representations</i> Adriane Boyd	41
<i>Usage of XSL Stylesheets for the Annotation of the Sámi Language Corpora.</i> Saara Huhmarniemi, Sjur N. Moshagen and Trond Trosterud	45
<i>Criteria for the Manual Grouping of Verb Senses</i> Cecily Jill Duffield, Jena D. Hwang, Susan Windisch Brown, Dmitriy Dligach, Sarah E. Vieweg, Jenny Davis and Martha Palmer	49
<i>Semi-Automated Named Entity Annotation</i> Kuzman Ganchev, Fernando Pereira, Mark Mandel, Steven Carroll and Peter White	53
<i>Querying Multimodal Annotation: A Concordancer for GeM</i> Martin Thomas	57
<i>Annotating Chinese Collocations with Multi Information</i> Ruifeng Xu, Qin Lu, Kam-Fai Wong and Wenjie Li	61
<i>Computing Translation Units and Quantifying Parallelism in Parallel Dependency Treebanks</i> Matthias Buch-Kromann	69
<i>Adding Semantic Role Annotation to a Corpus of Written Dutch</i> Paola Monachesi, Gerwert Stevens and Jantine Trapman	77
<i>A Search Tool for Parallel Treebanks</i> Martin Volk, Joakim Lundborg and Maël Mettler	85
<i>Annotating Expressions of Appraisal in English</i> Jonathon Read, David Hope and John Carroll	93

<i>Active Learning for Part-of-Speech Tagging: Accelerating Corpus Annotation</i>	
Eric Ringger, Peter McClanahan, Robbie Haertel, George Busby, Marc Carmen, James Carroll, Kevin Seppi and Deryle Lonsdale	101
<i>Combining Independent Syntactic and Semantic Annotation Schemes</i>	
Marc Verhagen, Amber Stubbs and James Pustejovsky	109
<i>XARA: An XML- and Rule-based Semantic Role Labeler</i>	
Gerwert Stevens	113
<i>ITU Treebank Annotation Tool</i>	
Gülşen Eryiğit	117
<i>Two Tools for Creating and Visualizing Sub-sentential Alignments of Parallel Text</i>	
Ulrich Germann	121
<i>Building Chinese Sense Annotated Corpus with the Help of Software Tools</i>	
Yunfang Wu, Peng Jin, Tao Guo and Shiwen Yu	125
<i>Annotating a Japanese Text Corpus with Predicate-Argument and Coreference Relations</i>	
Ryu Iida, Mamoru Komachi, Kentaro Inui and Yuji Matsumoto	132
<i>Web-based Annotation of Anaphoric Relations and Lexical Chains</i>	
Maik Stührenberg, Daniela Goecke, Nils Diewald, Alexander Mehler and Irene Cramer	140
<i>Standoff Coordination for Multi-Tool Annotation in a Dialogue Corpus</i>	
Kepa Joseba Rodríguez, Stefanie Dipper, Michael Götze, Massimo Poesio, Giuseppe Riccardi, Christian Raymond and Joanna Rąbiega-Wiśniewska	148
<i>PoCoS - Potsdam Coreference Scheme</i>	
Olga Krasavina and Christian Chiarcos	156
<i>Multiple-step Treebank Conversion: From Dependency to Penn Format</i>	
Cristina Bosco	164
<i>Experiments with an Annotation Scheme for a Knowledge-rich Noun Phrase Interpretation System</i>	
Roxana Girju	168
<i>IGT-XML: An XML Format for Interlinearized Glossed Text</i>	
Alexis Palmer and Katrin Erk	176
<i>Shared Corpora Working Group Report</i>	
Adam Meyers, Nancy Ide, Ludovic Denoyer and Yusuke Shinyama	184
<i>Discourse Annotation Working Group Report</i>	
Manfred Stede, Janyce Wiebe, Eva Hajicova, Brian Reese, Simone Teufel, Bonnie Webber and Theresa Wilson	191