<SRL>HPSG</>

Proceedings

of

The First Workshop on Treebanks and Linguistic Theories (TLT2002)

20th and 21st September 2002, Sozopol, Bulgaria

SfS, Tuebingen University, Tuebingen, Germany
LML, Bulgarian Academy of Sciences, Sofia, Bulgaria

Editors:

Erhard Hinrichs
Kiril Simov

Local Organisation:

Kiril Simov
Petya Osenova
Milena Slavcheva

Co-located with Fall School on Empirical Linguistics and Natural Language Processing, 2002

Workshop Organisers

Erhard Hinrichs
Seminar für Sprachwissenschaft
Universität Tübingen
Eh@sfs.uni-tuebingen.de

Kiril Simov
BulTreeBank Project
http://www.BulTreeBank.org
Linguistic Modelling Laboratory, CLPP,
Bulgarian Academy of Sciences
kivs@bultreebank.org

Workshop Programme Committee

Erhard Hinrichs, Germany (co-chair)
Tilman Berger , Germany
Marek Swidzinski, Poland
Adam Przepiórkowski, Poland
Kiril Simov, Bulgaria (co-chair)
Vladimir Petkevič, Czech Republic
Anatolij N. Baranov, Russia
Sandra Kübler, Germany
Kemal Oflazer, Turkey
Michael Barlow, USA
Tomaz Erjavec, Slovenia
Robert Engels, Norway
Andreas Wagner, Germany
Frank Richter, Germany
Manfred Sailer, Germany
Walter Daelemans, Belgium
Karel Oliva, Austria
Laurent Romary, France

Local Organisation

Kiril Simov, Petya Osenova, Milena Slavcheva

BulTreeBank Project
http://www.BulTreeBank.org
Linguistic Modelling Laboratory, CLPP,Bulgarian Academy of Sciences

Table of Contents

Elisaveta Balabanova and Krassimira Ivanova
Creating a machine-readable version of Bulgarian valence dictionary: (A case study of CLaRK system application) 1

Philippe Blanche and Marie-Laure Guénot
Flexible Corpus Annotation with Property Grammar 13

Sabine Brants, Stefanie Dipper, Silvia Hansen, Wolfgang Lezius, George Smith
TIGER Treebank 24

Aoife Cahill, Mairéad McCarthy, Josef Van Genabith and Andy Way
Evaluating Automatic F-Structure Annotation for the Penn-II Treebank 43

Montserrat Civit, M. Martí Antònín
Design Principles for a Spanish Treebank 61

Erhard Hinrichs and Julia Trushkina
Forging Agreement: Morphological Disambiguation of Noun Phrases 78

Krassimira Ivanova, Dimitar Doikoff
Cascaded Regular Grammars and Constraints over Morphologically Annotated Data for Ambiguity Resolution 96

Jiří Mírovský, Roman Ondruška and Daniel Průša
Searching through Prague Dependency Treebank Conception and Architecture 114

Joakim Nivre
What kinds of trees grow in Swedish soil? A Comparison of Four Annotation Schemes for Swedish 123

Stephan Oepen, Dan Flickinger, Kristina Toutanova, Christoper D. Manning
LinGO Redwoods: A Rich and Dynamic Treebank for HPSG 139

Petya Osenova
Bulgarian Nominal Chunks and Mapping Strategies for Deeper Syntactic Analyses 150

Petya Osenova and Sia Kolkovska
Combining the named-entity recognition task and NP chunking strategy for robust pre-processing 167

Kiril Simov, Alexander Simov, Milen Kouylekov, Krassimira Ivanova
CLaRK System: Construction of Treebanks 183

Milena Slavcheva
Segmentation Layers in the Group of the Predicate: a Case Study of Bulgarian within the BulTreeBank Framework 199

Oliver Streiter
Treebank Development with Deductive and Abductive Explanation-based Learning: Exploratory Experiments 211

Yovka Tisheva and Marina Dzhonova
Information Structure Level in TreeBanks 231

Kristina Toutanova, Christoper D. Manning, Stephan Oepen
Parse Ranking for a Rich HPSG Grammar 253

Tzvetomira Venkova
Bilingual corpora as a platform for cross-linguistic treebank development 264

Author Index

Elisaveta Balabanova	1
Philippe Blanche	13
Sabine Brants	24
Aoife Cahill	43
Montserrat Civit	61
Stefanie Dipper	24
Dimitar Doikoff	96
Marina Dzhonova	231
Dan Flickinger	139
Marie-Laure Guénot	13
Silvia Hansen	24
Erhard Hinrichs	78
Krassimira Ivanova	1,96,183
Sia Kolkovska	167
Milen Kouylekov	183
Wolfgang Lezius	24
Christoper D. Manning	139, 253
M. Martí Antònín	61
Mairéad McCarthy	43
Jiří Mírovský	114
Joakim Nivre	123
Stephan Oepen	139, 253
Roman Ondruška	114
Petya Osenova	150, 167
Daniel Průša	114
Alexander Simov	183
Kiril Simov	183
Milena Slavcheva	199
George Smith	24
Oliver Streiter	211
Yovka Tisheva	231
Kristina Toutanova	139, 253
Julia Trushkina	78
Josef Van Genabith	43
Tzvetomira Venkova	264
Andy Way	43