

EACL 2012

EACL 2012 Joint Workshop of LINGVIS & UNCLH
Visualization of Linguistic Patterns
and
Uncovering Language History from Multilingual Resources

Proceedings of the Workshop

April 23 - 24 2012
Avignon France

© 2012 The Association for Computational Linguistics

ISBN 978-1-937284-19-0

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Organizers:

LINGVIS:

Miriam Butt (University of Konstanz)
Sheelagh Carpendale (University of Calgary)
Gerald Penn (University of Toronto)

UNCLH:

Jelena Prokić (LMU Munich)
Michael Cysouw (LMU Munich)
Thomas Mayer (LMU Munich)
Steven Moran (LMU Munich)

Program Committee:

Quentin Atkinson (University of Auckland)
Christopher Collins (University of Ontario)
Chris Culy (University of Tübingen)
Dan Dediu (MPI Nijmegen)
Michael Dunn (MPI Nijmegen)
Sheila Embleton (York University, Toronto)
Simon Greenhill (University of Auckland)
Harald Hammarström (University of Nijmegen)
Annette Hautli (University of Konstanz)
Wilbert Heeringa (Meertens Institute, Amsterdam)
Gerhard Heyer (University of Leipzig)
Eric Holman (UCLA)
Gerhard Jäger (University of Tübingen)
Daniel Keim (University of Konstanz)
Tibor Kiss (University of Bochum)
Jonas Kuhn (University of Stuttgart)
John Nerbonne (University of Groningen)
Anke Lüdeling (Humboldt University, Berlin)
Don Ringe (University of Pennsylvania)
Christian Rohrdantz (University of Konstanz)
Tandy Warnow (University of Texas at Austin)
Søren Wichmann (EVA MPI, Leipzig)

Invited Speakers:

Daniela Oelke (University of Konstanz)
Grzegorz Kondrak (University of Alberta)

Table of Contents

| | |
|---|-----|
| <i>Introduction</i> | |
| Miriam Butt, Jelena Prokić, Thomas Mayer and Michael Cysouw | 1 |
| <i>Lexical Semantics and Distribution of Suffixes - A Visual Analysis</i> | |
| Christian Rohrdantz, Andreas Niekler, Annette Hautli, Miriam Butt and Daniel A. Keim | 7 |
| <i>Looking at word meaning. An interactive visualization of Semantic Vector Spaces for Dutch synsets</i> | |
| Kris Heylen, Dirk Speelman and Dirk Geeraerts | 16 |
| <i>First steps in checking and comparing Princeton WordNet and Estonian Wordnet</i> | |
| Ahti Lohk, Kadri Vare and Leo Võhandu | 25 |
| <i>Visualising Typological Relationships: Plotting WALS with Heat Maps</i> | |
| Richard Littauer, Rory Turnbull and Alexis Palmer | 30 |
| <i>Automating Second Language Acquisition Research: Integrating Information Visualisation and Machine Learning</i> | |
| Helen Yannakoudakis, Ted Briscoe and Theodora Alexopoulou | 35 |
| <i>Visualising Linguistic Evolution in Academic Discourse</i> | |
| Verena Lyding, Ekaterina Lapshinova-Koltunski, Stefania Degaetano-Ortlieb, Henrik Dittmann and Chris Culy | 44 |
| <i>Similarity Patterns in Words (Invited talk)</i> | |
| Grzegorz Kondrak | 49 |
| <i>Language comparison through sparse multilingual word alignment</i> | |
| Thomas Mayer and Michael Cysouw | 54 |
| <i>Recovering dialect geography from an unaligned comparable corpus</i> | |
| Yves Scherrer | 63 |
| <i>Detecting Shibboleths</i> | |
| Jelena Prokić, Çağrı Cöltekin and John Nerbonne | 72 |
| <i>Estimating and visualizing language similarities using weighted alignment and force-directed graph layout</i> | |
| Gerhard Jäger | 81 |
| <i>Explorations in creole research with phylogenetic tools</i> | |
| Aymeric Daval-Markussen and Peter Bakker | 89 |
| <i>Tracking the dynamics of kinship and social category terms with AustKin II</i> | |
| Patrick McConvell and Laurent Dousset | 98 |
| <i>Using context and phonetic features in models of etymological sound change</i> | |
| Hannes Wettig, Kirill Reshetnikov and Roman Yangarber | 108 |
| <i>LexStat: Automatic Detection of Cognates in Multilingual Wordlists</i> | |
| Johann-Mattis List | 117 |

Conference Program

Introduction

Miriam Butt, Jelena Prokić, Thomas Mayer and Michael Cysouw

Monday, April 23, 2012

- 9:00 Invited talk by Christopher Collins
- 10:00 Coffee break
- 10:30 *Lexical Semantics and Distribution of Suffixes - A Visual Analysis*
Christian Rohrdantz, Andreas Niekler, Annette Hautli, Miriam Butt and Daniel A. Keim
- 11:15 *Looking at word meaning. An interactive visualization of Semantic Vector Spaces for Dutch synsets*
Kris Heylen, Dirk Speelman and Dirk Geeraerts
- 12:00 *First steps in checking and comparing Princeton WordNet and Estonian Wordnet*
Ahti Lohk, Kadri Vare and Leo Võhandu
- 12:30 Lunch break
- 14:30 *Visualising Typological Relationships: Plotting WALs with Heat Maps*
Richard Littauer, Rory Turnbull and Alexis Palmer
- 15:00 *Automating Second Language Acquisition Research: Integrating Information Visualisation and Machine Learning*
Helen Yannakoudakis, Ted Briscoe and Theodora Alexopoulou
- 15:30 *Visualising Linguistic Evolution in Academic Discourse*
Verena Lyding, Ekaterina Lapshinova-Koltunski, Stefania Degaetano-Ortlieb, Henrik Dittmann and Chris Culy
- 16:15 *The Potential of Visual Analytics for Computational Linguistics (Invited talk)*
Daniela Oelke
- 17:15 Discussion

Tuesday, April 24, 2012

- 9:00 *Similarity Patterns in Words (Invited talk)*
Grzegorz Kondrak
- 10:30 Coffee break
- 10:30 *Language comparison through sparse multilingual word alignment*
Thomas Mayer and Michael Cysouw
- 11:00 *Recovering dialect geography from an unaligned comparable corpus*
Yves Scherrer
- 11:30 *Detecting Shibboleths*
Jelena Prokić, Çağrı Cöltekin and John Nerbonne
- 12:00 *Estimating and visualizing language similarities using weighted alignment and force-directed graph layout*
Gerhard Jäger
- 12:30 Lunch break
- 14:30 *Explorations in creole research with phylogenetic tools*
Aymeric Daval-Markussen and Peter Bakker
- 15:00 *Tracking the dynamics of kinship and social category terms with AustKin II*
Patrick McConvell and Laurent Dousset
- 15:30 Coffee break
- 16:00 *Using context and phonetic features in models of etymological sound change*
Hannes Wettig, Kirill Reshetnikov and Roman Yangarber
- 16:30 *LexStat: Automatic Detection of Cognates in Multilingual Wordlists*
Johann-Mattis List
- 17:00 Discussion