



2013

Programme

NODALIDA

22 – 24 May, 2013 • Oslo

Wednesday, 22 May

08:30 – 09:00	Registration	Lobby
11:00 – 15:00	Registration	Lobby
09:00 – 17:30	NODALIDA Workshops	
	Second Workshop on NLP for Computer-Assisted Language Learning Afternoon workshop, starting with lunch at 12.00	Grupperom 1
	Workshop on Computational Historical Linguistics Whole day workshop	Undervisningsrom 1
	Workshop on Lexical Semantic Resources for NLP Afternoon workshop, starting with lunch at 12.00	Undervisningsrom 2
	Workshop on Nordic Language Research Infrastructure Afternoon workshop, starting with lunch at 12.00	Undervisningsrom 3
19:30	Conference Dinner Restaurant Argent, Oslo Opera House Please bring your drink code ticket and place it next to your glass at the table. You will also need it to claim your aperitif.	

Thursday, 23 May

08.30– 09.00	Registration	Lobby
09:00– 09:20	Opening <i>Inga Bostad</i> (Prorector of the University of Oslo) <i>Krister Lindén</i> (President of NEALT) <i>Stephan Oepen</i> (Chair of the Program Committee) <i>Janne Bondi Johannessen</i> (Chair of the local Organizing Committee)	Aud 1
09:20– 10:20	Keynote (Chair: Lars Ahrenberg) The Conversational User Interface <i>Ron Kaplan</i>	Aud 1
10:20– 10:40	Coffee Break	Basement
10:40– 12:40	Oral Presentations Parallel Session 1 (Chair: Janne Bondi Johannessen) Tagging the Past: Experiments using the Saga Corpus <i>Hrafn Loftsson</i> Normalisation of Historical Text Using Context-Sensitive Weighted Levenshtein Distance and Compound Splitting <i>Eva Pettersson, Beáta Megyesi, Joakim Nivre</i> Using Constraint Grammar for Chunking <i>Eckhard Bick</i> Using Finite State Transducers for Making Efficient Reading Comprehension Dictionaries	Aud 1

Thursday, 23 May

10:40–
12:40

Parallel Session 2 (Chair: Kimmo Koskenniemi)

Aud 2

Towards a Dependency-Based PropBank of General Finnish
Katri Haverinen, Veronika Laippala, Samuel Kohonen, Anna Missilä, Jenna Nyblom, Stina Ojala, Timo Viljanen, Tapio Salakoski, Filip Ginter

Classifying Multimodal Turn Management in Danish Dyadic First Encounters
Costanza Navarretta, Patrizia Paggio

Nordic and Baltic Wordnets Aligned and Compared through "WordTies"

Bolette S. Pedersen, Lars Borin, Markus Forsberg, Neeme Kahusk, Krister Lindén, Jyrki Niemi, Niklas Nisbeth, Lars Nygaard, Heili Orav, Eiríkur Rögnvaldsson, Mitchell Seaton, Kadri Vider, Kaarlo Voionmaa

Baltic and Nordic Parts of the European Linguistic Infrastructure

Inguna Skadiņa, Andrejs Vasiljevs, Lars Borin, Krister Lindén, Gyri Losnegaard, Sussi Olsen, Bolette S. Pedersen, Roberts Rozis, Koenraad De Smedt

12:40–
13:40

Lunch

13:40–
14:40

Keynote (Chair: Bolette S. Pedersen)

Detecting and Processing Figurative Language in Discourse
Caroline Sporleder

Thursday, 23 May

15:50– **Coffee** Basement
16:10

16:10– **Special Session on HPC for NLP** Aud 1
17:00 (Chair: Koenraad de Smedt)

Experiences in Building the Let's MT! Portal on
Amazon EC2
Jörg Tiedemann

The Nordic e-Infrastructure Initiative: Opportunities for
Cross-Border Collaboration
Gudmund Høst

Tidying up the Basement: A Tale of Large-Scale
Parsing on National eInfrastructure
Stephan Oepen

18:00 **Reception at the Oslo City Hall with Standing Buffet**
NOTE: Dress code: Smart casual. Please bring your
invitation and be there at 17.50 at the latest.

Thursday, 23 May

14:40–
15:50

Posters & Demonstrations

Basement

Building a Large Automatically Parsed Corpus of Finnish
Filip Ginter, Jenna Nyblom, Veronika Laippala, Samuel Kohonen, Katri Haverinen, Simo Vihjanen, Tapio Salakoski

New Measures to Investigate Term Typology by
Distributional Data
Jussi Karlgren

Analysis of Phonetic Transcription for Danish Automatic
Speech Recognition
Andreas Søeborg Kirkedal

Combining Statistical Machine Translation and
Translation Memories with Domain Adaptation
Samuel Lübbli, Mark Fishel, Martin Volk, Manuela Weibel

Simple and Accountable Segmentation of Marked-up
Text
Jonathon Read, Rebecca Dridan, Stephan Oepen

The Automatic Identification of Discourse Units in Dutch
Text
Nynke Van Der Vliet, Gosse Bouma, Gisela Redeker

Example-Based Treebank Querying with GrETEL - Now
Also for Spoken Dutch
Liesbeth Augustinus, Vincent Vandeghinste, Ineke Schuurman, Frank Van Eynde

Korp and Karp – A Bestiary of Language Resources: The
Research Infrastructure of Språkbanken
Malin Ahlberg, Lars Borin, Markus Forsberg, Martin Hammarstedt, Leif-Jöran Olsson, Olof Olsson, Johan Roxendal, Jonatan Uppström

A Grammar Sparrer for Norwegian
Lars Hellan, Tore Bruland, Elias Aamot, Mads H. Sandøy

Finite State Applications with Javascript
Mans Hulden, Miikka Silfverberg, Jerid Francom

Friday, 24 May

09:00–
10:00

Oral Presentations

Parallel Session 1 (Chair: Victoria Rosén)

Aud 1

Features Indicating Readability in Swedish Text
*Johan Falkenjack, Katarina Heimann Mühlenbock,
Arne Jönsson*

Tagging a Morphologically Complex Language
Using an Averaged Perceptron Tagger: The Case of
Icelandic
Hrafn Loftsson, Robert Östling

Parallel Session 2 (Chair: Costanza Navarretta)

Aud 2

Exploring Features for Named Entity Recognition
in Lithuanian Text Corpus
*Jurgita Kapočiūtė-Dzikienė, Anders Nøklestad,
Janne Bondi Johannessen, Algis Krupavičius*

IPhraxtor: A Linguistically Informed System for
Extraction of Term Candidates
Magnus Merkel, Jody Foo, Lars Ahrenberg

Friday, 24 May

10:00–
10:20

Lightning Talks

Aud 1

Parallel Session 1 (Chair: Victoria Rosén)

From Speech Corpus to Intonation Corpus:
Clustering Phrase Pitch Contours of Lithuanian
Gailius Raškinis, Asta Kazlauskienė

Statistical Machine Translation with Readability
Constraints
*Sara Stymne, Jörg Tiedemann, Christian Hardmeier,
Joakim Nivre*

Parallel Session 2 (Chair: Costanza Navarretta)

Aud 2

Negation Scope Delimitation in Clinical Text Using
Three Approaches: NegEx, PyConTextNLP and
SynNeg
*Hideyuki Tanushi, Hercules Dalianis, Martin Duneld,
Maria Kvist, Maria Skeppstedt, Sumithra Velupillai*

Tone Restoration in Transcribed Kammu:
Decision-List Word Sense Disambiguation for an
Unwritten Language
Marcus Uneson

Friday, 24 May

10:20–

Posters & Demonstrations

Basement

11:20

From Speech Corpus to Intonation Corpus: Clustering Phrase
Pitch Contours of Lithuanian

Gailius Raškinis, Asta Kazlauskienė

Statistical Machine Translation with Readability Constraints

*Sara Stymne, Jörg Tiedemann, Christian Hardmeier, Joakim
Nivre*

Negation Scope Delimitation in Clinical Text Using Three
Approaches: NegEx, PyConTextNLP and SynNeg

*Hideyuki Tanushi, Hercules Dalianis, Martin Duneld, Maria
Kvist, Maria Skeppstedt, Sumithra Velupillai*

Tone Restoration in Transcribed Kammu: Decision-List Word
Sense Disambiguation for an Unwritten Language

Marcus Uneson

Modeling OOV Words With Letter N-Grams in Statistical
Taggers: Preliminary Work in Biomedical Entity Recognition

Teemu Ruokolainen, Miikka Silfverberg

Constructing a Multilingual Database of Verb Valence

Lars Hellan, Tore Bruland

Building an Open-Source Development Infrastructure for
Language Technology Projects

Sjur N. Moshagen, Tommi A. Pirinen, Trond Trosterud

The INESS Treebanking Infrastructure

*Paul Meurer, Helge Dyvik, Victoria Rosén, Koenraad De Smedt,
Gunn Inger Lyse, Gyri Smørdal Losnegaard, Martha Thunes*

HPC-ready Language Analysis for Human Beings

*Emanuele Lapponi, Erik Velldal, Nikolay A. Vazov, Stephan
Oepen*

Building Gold-Standard Treebanks for Norwegian

Per Erik Solberg

Friday, 24 May

11:20–11:40	Coffee	Basement
11:40–12:30	Business Meeting	Aud 1
12:30–13:30	Lunch	
13:30–15:00	Student Session: Oral Presentations	
	Parallel Session 1 (Chair: Joerg Tiedemann)	Aud 1
	The IPP Effect in Afrikaans: A Corpus Analysis <i>Liesbeth Augustinus, Peter Dirix</i>	
	Using Factual Density to Measure Informativeness of Web Documents <i>Christopher Horn, Alisa Zhila, Alexander Gelbukh, Roman Kern, Elisabeth Lex</i>	
	Using Topic Models in Content-Based News Recommender Systems <i>Tapio Luostarinen, Oskar Kohonen</i>	
	Parallel Session 2 (Chair: Heiki-Jaan Kaalep)	Aud 2
	Bootstrapping an Unsupervised Approach for Classifying Agreement and Disagreement <i>Bernd Opitz, Căcilia Zirn</i>	
	Morphological Analysis with Limited Resources: Latvian Example <i>Pēteris Paikens, Laura Rituma, Lauma Pretkalniņa</i>	
	Statistical Syntactic Parsing for Latvian <i>Lauma Pretkalniņa, Laura Rituma</i>	
15:00–16:00	Keynote (Chair: Stephan Oepen)	Aud 1
	6,909 Reasons to Mess Up Your Data <i>Anders Søgaard</i>	
16:00–16:10	Closing	Aud 1
16:10–16:30	Coffee	Basement

Info

Venue

Georg Sverdrups hus Blindern, University of Oslo

Lunch

Frederikke Building, across the square above Georg Sverdrups hus

WiFi

Network: conferences

Password: uio202aar

Eduroam is also available.

Sponsors

Conference sponsors

WeSearch

iness

Lingit

max manus

Forskningsrådet

LETTERSTEDTSKA
FÖRENINGEN

NEALT sponsors

computas

The Center of Estonian
Language Resources

DET HUMANISTISKE FAKULTET
KØBENHAVNS UNIVERSITET

GSLT

Lingsoft
LANGUAGE
SOLUTIONS



Mikro Værkstedet



National Library of Norway

textUrgy

