

NEALT PROCEEDINGS SERIES

VOL. #

*Nordic Perspectives on the CLARIN Infrastructure
of Common Language Resources
– Workshop Proceedings*

May 14, 2009

Odense, Denmark

Nodalida 2009

Editors

Rickard Domeij
Kimmo Koskenniemi
Steven Krauwer
Bente Maegaard
Eiríkur Rögnvaldsson
Koenraad de Smedt

NORTHERN EUROPEAN ASSOCIATION FOR LANGUAGE TECHNOLOGY

Nordic Perspectives on the CLARIN Infrastructure of Language Resources
– Workshop Proceedings
NEALT Proceedings Series, Vol. #

© 2009 The editors and contributors.

ISSN 1736-6305

Published by

Northern European Association for Language
Technology (NEALT)
<http://omilia.uio.no/nealt>

Electronically published at

Tartu University Library (Estonia)
<http://dspace.utlib.ee/dspace/handle/10062/4116>

Volume Editors

Rickard Domeij
Kimmo Koskenniemi
Steven Krauwer
Bente Maegaard
Eiríkur Rögnvaldsson
Koenraad de Smedt

Series Editor-in-Chief

Mare Koit

Series Editorial Board

Lars Ahrenberg
Koenraad De Smedt
Kristiina Jokinen
Joakim Nivre
Patrizia Paggio
Vytautas Rudzionis

Contents

Preface	iv
Maia Andréasson, Lars Borin, Markus Forsberg, Jonas Beskow, Rolf Carlson, Jens Edlund, Kjell Elenius, Kahl Hellmer, David House, Magnus Merkel, Eva Forsbom, Beáta Megyesi, Anders Eriksson & Sven Strömqvist: <i>Swedish CLARIN activities</i>	1
Hanne Fersøe & Bente Maegaard: <i>CLARIN in Denmark – European and Nordic perspectives</i>	6
Kimmo Koskenniemi & Antii Arppe: <i>Nordic co-operation in building the language resource infrastructures</i>	9
Rūta Marcinkevičienė: <i>Two decades of Lithuanian HLT</i>	13
Einar Meister, Tiit Roosmaa & Jaak Vilo: <i>Estonian language technology Anno 2009</i>	17
Eiríkur Rögnvaldsson, Hrafn Loftsson, Kristín Bjarnadóttir, Marhew Whelpton, Anna Björk Nikulásdóttir & Anton Karl Ingason: <i>Icelandic Language Resources and Technology: Status and Prospects</i>	23
Inguna Skandiņa: <i>CLARIN in Latvia: current situation and future perspectives</i>	29
Pavel Skrelin, Vera Evdokimova & Karina Evgrafova: <i>The Possible NEALT Role in the Consolidation of the Nordic and Baltic Language Resources</i>	34
Koenraad de Smedt: <i>CLARIN: Norwegian and Nordic perspectives</i>	39

Preface

The Nordic countries have a long tradition of cooperating within many areas, including politics, education and science. Many languages are closely related and sometimes also the same language is spoken over national boundaries (for example Sámi and Swedish). Language technology is relatively well developed in these countries, but much more is needed to build the infrastructure needed for advanced R&D, and to secure the languages of the region for the future. The CLARIN project is an initiative on the European level to meet those challenges by making language resources and technology available and usable.

In recent years, new regions around the Baltic have become parts of the Nordic area. With increased cooperation, coordination and consolidation of common strengths, the Nordic/Baltic countries could strengthen their work in language technology infrastructure and their contribution to CLARIN.

The main topic of the workshop is Nordic strengths and opportunities of cooperation within the NEALT Geographic Region in constructing an infrastructure of common language resources in connection to the European CLARIN (Common Language Resources and Technology Infrastructure) initiative.

The purpose is to find ways of cooperating that will strengthen the contribution of the associated countries to the development of an infrastructure of common language resources within the CLARIN initiative. In the workshop, participants in CLARIN from the NEALT-associated countries will be given the chance to present their national projects and discuss possible ways of cooperating, sharing resources, coordinating activities, consider new projects and such. Opportunities and proposals for closer cooperation and coordination will be presented and discussed at the workshop.

CLARIN participants in the NEALT-associated countries were invited to present their national work from the perspective of possible cooperation between groups and projects in the different countries: Denmark, Estonia, Finland, Iceland, Latvia, Lithuania, Norway and Sweden. The workshop is intended for participants having an interest in developing the language resources and technology in the NEALT Geographic Region for languages spoken in that region. There will be an opportunity for each such country to present an overview of the status of their national language resource infrastructure.

Eight abstracts from each of the above mentioned countries were submitted for review by the editors (who did not review contributions from their own country). All of them were accepted. We, the editors, want to thank the authors for their contributions. We look forward to a promising workshop in Odense where they will be presented and discussed.

Rickard Domeij, Kimmo Koskenniemi, Steven Krauwer,
Bente Maegaard, Eiríkur Rögnvaldsson & Koenraad De Smedt

*Nordic Perspectives on the CLARIN
Infrastructure
of Common Language Resources
– Workshop Proceedings*

May 14, 2009

Odense, Denmark

Nodalida 2009

Workshop website:

<https://kitwiki.csc.fi/twiki/bin/view/Nealt/SigInfraNodalida2009WorkshopCall>

Conference website:

<http://beta.visl.sdu.dk/nodalida2009/>

Organizers

Rickard Domeij, the Language Council of Sweden, chairman of the Nordic language councils' working group on language technology and chairman of the NEALT [SigInfra](#)

Kimmo Koskenniemi, Prof. of language technology at the University of Helsinki, Department of General Linguistics, Executive Board member of CLARIN

Steven Krauwer, Utrecht University, CLARIN Coordinator

Bente Maegaard, Head of Centre for Language Technology, University of Copenhagen, Denmark. Executive Board member of CLARIN

Eiríkur Rögnvaldsson, Prof. of Icelandic Language, University of Iceland

Koenraad De Smedt, Prof. of Computational Linguistics, University of Bergen; national coordinator of CLARIN in Norway

