NODALIDA '99

Proceedings from the 12th "Nordiske datalingvistikkdager".

Trondheim, 9-10 December, 1999.

Torbjørn Nordgård, editor

Trondheim

Department of Linguistics, NTNU

2000

This volume and the conference were sponsored by:
Telenor
Faculty of Arts, NTNU
Department of Linguistics, NTNU

Published by
Department of Linguistics
Norwegian University of Science and Technology
N-7491 Trondheim, Norway

Printed by
NTNU

ISBN 82-92225-00-5

Contents

Tore Amble
BusTUC – A natural language bus route adviser in Prolog 1

Antti Arppe
Developing a grammar checker for Swedish 13

Juhani Birn
Detecting grammar errors with Lingsoft's Swedish grammar checker 28

Lars Borin
Pivot Alignment 41

Rickard Domeij, Ola Knutsson, Johan Carlberger & Viggo Kann
Granska – an efficient hybrid system for Swedish grammar checking 49

Kristofer Franzén
Adapting an English Information Extraction System to Swedish 57

Kristin Hagen, Janne Bondi Johannessen & Anders Nøklestad
The shortcomings of a tagger 66

Anette Hulth & Lars Asker
Merging Classifiers for Improved Information Retrieval 76

Anna Jonsson
Extracting Keywords from Digital Document Collections 83

Atanas K. Kiryakov & Kiril Iv. Simov
Ontologically Supported Semantic Matching 91

Janne Lindberg
Automatic Detecting of Lexicalised Phrases in Swedish 103

Beáta Megyesi & Sara Rydin
Towards a Finite-State Parser for Swedish 115

Costanza Navaretta
Semantic Clustering of Adjectives and Verbs Based on Syntactic Patterns 124

Anne Neville
An HPSG Account of Danish Pre-nominals 134

Arild Noven, Per Arne Larsen, Bente Moxness & Kolbjørn Slethei
Tonem 1 eller 2 eller 1,5? 143

Patrizia Paggio
Syntactic Analysis and Error Correction for Danish in the SCARRIE Project 152

Botond Pakucs & Björn Gambäck
Designing a System for Swedish Spoken Document Retrieval 162

Stina Nylander
Statistics and Phonotactical Rules in Finding OCR Errors 174

Paulo Quaresma & Irene Pimenta Rodrigues
An Information Retrieval System with Co-operative Behaviour 183

Diana Santos & Signe Oksefjell
An Evaluation of the Translation Corpus Aligner, with special reference to the language pair English-Portuguese 191

Koenraad de Smedt & Victoria Rosén
Automatic proofreading for Norwegian: The challenges of lexical and grammatical variation 206

Jörg Tiedeman
Word Alignment Step by Step 216

Heli Uibo
On Using the Two-level Model as the Basis of Morphological Analysis and Synthesis of Estonian 228

Andy Way
LFG-DOT: Combining Constraint-Based and Empirical Methodologies for Robust MT 243