

**COLING - The 20th International Conference on
Computational Linguistics**

Workshop

**eLearning for Computational Linguistics and
Computational Linguistics for eLearning**

Geneva, 28. August 2004

Proceedings

Editors: Lothar Lemnitzer, Detmar Meurers, Erhard Hinrichs

Introduction

The idea of this workshop arose on the occasion of an international workshop on *eLearning and Computational Linguistics*, organized by the MiLCA project and held in Blaubeuren in November 2003 (in the MiLCA project - Medienintensive Lehrmodule in der Computerlinguistik-Ausbildung, cf. <http://milca.sfs.uni-tuebingen.de> -, five German partners are developing core modules for a CL curriculum as web based trainings or virtual courses.). It, therefore, seemed fitting and timely to identify this topic as one of the leading themes of the workshop. The COLING 2004 conference appeared to the participants as a good opportunity for such an event. We are very grateful to the COLING organizers to have given us the opportunity to conduct a workshop with the title *eLearning for Computational Linguistics and Computational Linguistics for eLearning*. This title reflects the two main topics that the participants of the MiLCA workshop had identified as the two leading themes to be addressed.

eLearning for Computational Linguistics

MiLCA is not the only project focussing on eLearning in support of teaching Computational Linguistics. We are aware of numerous individual and institutional efforts of the kind under way. This topic should also be viewed as a follow-up of the 1999 EACL workshop *Computer and Internet Supported Education in Language and Speech Technology*, the 2002 ACL workshop on *Effective Tools and Methodologies for Teaching NLP*, and the LREC 2004 workshop on *Language Resources: Integration and Development in eLearning and in Teaching Computational Linguistics*. It will give practioners in the field the opportunity to share didactic experience of eTeaching CL, to identify and to disseminate good practice models and resources and to share novel approaches to eLearning in CL.

Computational Linguistics for eLearning

This part of the workshop addresses the role which language technology plays in the development of intelligent platforms for eLearning. One focus of this part of the workshop concerns the contribution of NLP to the building of intelligent computer-aided language learning (I-CALL) systems. Another leading question of this topic area concerns the way in which NLP can improve the performance of eLearning environments in general.

We were glad to receive a large number of high-quality papers, most of which will be presented on the workshop and can be found in this volume.

For presentation on the workshop, we have grouped the papers into the following sections:

- The first paper gives an overview of the issues involved in NLP for eLearning. Some approaches are outlined and illustrated by projects of the author and his colleagues;
- This is followed by three papers which present examples of language learning systems and environments. These papers focus on the role which NLP plays in them;
- The largest group of five papers presents ways in which NLP can help teachers to assess and evaluate the response of students to tasks of various kinds;
- Two papers describe computational systems which help students of linguistics and computational linguistics to acquire relevant knowledge by e-learning;
- The last three papers present further important aspects of NLP and eLearning: in the first paper, an NLP-driven environment for foreign language learning in primary schools is presented; the second paper focuses on the use of semi-automatic annotation at different stages of the learning process; the last paper deals with the use of Text Mining in the context of learning ancient Chinese poetry.

We hope that you will enjoy the workshop and the proceedings as much as we have enjoyed its preparation. May this workshop be a forum for the exchange of ideas on the intersection of Computational Linguistics and eLearning and a starting point for further fruitful co-operations.

The work of the Tübingen group in the preparation of this workshop has been partially supported by the German Federal Ministry of Education and Research (BMBF), in the context of the MiLCA project. We would like to express our gratitude for this support.

The organizers

Workshop chairs and programme committee

Lars Borin, University of Göteborg, Sweden

Kai-Uwe Carstensen, University of Zürich, Switzerland

Trude Heift, Simon Fraser University, Vancouver, Canada

Erhard Hinrichs, University of Tübingen, Germany, co-chair

Claudia Kunze, University of Tübingen, Germany

Lothar Lemnitzer, University of Tübingen, Germany co-chair

Petra Ludewig, University of Osnabrück, Germany

Detmar Meurers, Ohio State University Columbus, USA, co-chair

Wolf Paprotté, University of Münster, Germany

Veit Reuer, University of Osnabrück, Germany

Kiril Simov, Academy of Sciences, Sofia, Bulgaria

Pavel Smrž, Masaryk University, Brno, Czech Republic

Maik Stührenberg, University of Gießen, Germany

Tzvetomira Venkova, Sofia University St Klement Ohridski, Bulgaria

Maria Wolters, Rhetorical Systems, Edinburgh, UK

Workshop chairs and programme committee

Lars Borin, University of Göteborg, Sweden

Kai-Uwe Carstensen, University of Zürich, Switzerland

Trude Heift, Simon Fraser University, Vancouver, Canada

Erhard Hinrichs, University of Tübingen, Germany, co-chair

Claudia Kunze, University of Tübingen, Germany

Lothar Lemnitzer, University of Tübingen, Germany co-chair

Petra Ludewig, University of Osnabrück, Germany

Detmar Meurers, Ohio State University Columbus, USA, co-chair

Wolf Paprotté, University of Münster, Germany

Veit Reuer, University of Osnabrück, Germany

Kiril Simov, Academy of Sciences, Sofia, Bulgaria

Pavel Smrž, Masaryk University, Brno, Czech Republic

Maik Stührenberg, University of Gießen, Germany

Tzvetomira Venkova, Sofia University St Klement Ohridski, Bulgaria

Maria Wolters, Rhetorical Systems, Edinburgh, UK

Table of Contents

Integrating Natural Language Processing into E-Learning - A Case of Czech <i>Pavel Smrž</i>	1
NLP serving the cause of language learning <i>Frédérique Segond, Thibault Parmentier</i>	11
NLP-based scripting for CALL activities <i>G. Antoniadis, S. Echinard, O. Kraif, T. Lebarbé, M. Loiseau, C. Ponton</i>	18
Generated Narratives for Computer-aided Language Teaching <i>Michael Levison, Greg Lessard</i>	26
Indexing Student Essays Paragraphs Using LSA Over an Integrated Ontological Space <i>Gaston G. Burek, Maria Vargas-Vera, Emanuela Moreale</i>	32
E-Assessment using Latent Semantic Analysis in the Computer Science Domain: A Pilot Study <i>Pete Thomas, Debra Haley, Anne deRoeck, Marian Petre</i>	38
Towards deeper understanding and personalisation in CALL <i>Galia Angelova, Alben Strupchanska, Ognyan Kalaydijev, Milena Yankova, Svetla Boytcheva, Irena Vitanova, Preslav Nakov</i>	45
Automatic Measuring of English Language Proficiency using MT Evaluation Technology <i>Keiji Yasuda, Fumiaki Sugaya, Eiichiro Sumita, Toshiyuki Takezawa, Genichiro Kikui, Seiichi Yamamoto</i>	53
Sentence Completion Tests for Training and Assessment in a Computational Linguistics Curriculum <i>Cerstin Mahlow, Michael Hess</i>	61
The Syntax Student's Companion: an eLearning Tool designed for (Computational) Linguistics Students <i>Aurélien Max</i>	71
CL for CALL in the Primary School <i>Katrina Keogh, Thomas Koller, Monica Ward, Elaine Uí Dhonnchadha, Josef van Genabith</i>	79
A New E-Learning Paradigm through Annotating Operations <i>Hiroaki Saito, Kyoko Ohara, Kengo Sato, Kazunari Ito, Hiroyuki Okamoto</i>	86
From Text to Exhibitions: A New Approach for E-Learning on Language and Literature based on Text Mining <i>Qiaozhu Mei, Junfeng Hu</i>	90