

Human Language Technology Resources for Central European Languages: European Integration Issues

Panel Session at LREC2000

Zygmunt Vetulani

Adam Mickiewicz University
Department of Computer Linguistics and Artificial Intelligence
ul. Matejki 48/49, PL-60769 Poznań, Poland
<http://main.amu.edu.pl/~vetulani>
vetulani@amu.edu.pl

Abstract

Below, the reader will find a number of problems submitted to discussion at LREC2000 within the panel about "Human Language Technology Resources for Central European Languages: European integration issues". The post-panel report containing comments and conclusions will be available at the WEB site: <http://main.amu.edu.pl/~zlisi/news/lrec2000.htm>

1. Objectives of the Panel

The panel on European integration and Human Language Technology (HLT) resources is an invitation to reflect and discuss on the state of the art and future developments in the domain of Central European language resources. In the R&D domain this process seems to be moving in the right direction due to the formal opening of the 5th Framework Programme (FP5) to the EC countries. Is this really the case? Is the opening more than just a political declaration of intent? What is the situation in the domain of HLT resources? These questions need to be answered urgently. At this moment EU and CE countries have much to lose or to gain together. It is time to anticipate further needs in the HLT field. This panel is an initiative to encourage the vigilance, active co-operation, and co-ordination of the HLT resources, which will be essential prerequisites for integration in the near future.

2. Some Already Identified Problems

2.1. Technological Gap

How large is the technological gap between EU and CE countries in the field of Human Language Technology resources? Are there any special measures and priorities necessary to stop this gap? What are risks to see this gap growing in the future? How much co-ordination is there between research project funding programs and educational programs?

2.2. EU Integration

Political and awareness issues: What are the appropriate promotion measures for increasing awareness about information technologies in the EU-candidate countries? Since the role of language resources seems to be underestimated in the CE countries, how can lessons learned from promoting standardisation and resource creation in the EU be applied to the candidate countries?

2.3. Resources

What could be the role of ELRA (and similar

organisations) in the development of Human Language Technologies in the CE countries? Why are CE languages practically non-existent in the ELRA resource catalogue at present? How can this be improved in the efficient creation of new HLT resources for these countries in the context of IT resources in general?

2.4. Under-representation

Under-representation of CE researchers in the international bodies responsible for language resources.

2.5. Fifth Framework Programme

Does the Fifth Framework Programme meet needs and expectations connected with the development of basic language resources for CE languages? Is the FP5 sufficiently open to specific problems of EU candidate countries in the field of language resources?

2.6. Advanced Education

What is the potential for funding joint focussed advanced education in the EU and the candidate countries in the area of Human Language Technologies and related fields, in co-operation with existing exchange and summer school programmes?

3. Affiliation of Panelists

Chair/moderator:

Zygmunt Vetulani
Adam Mickiewicz University, Department of Computer Linguistics and Artificial Intelligence, Poznań, Poland

Participants:

Roberto Cencioni
European Commission, DG XIII, Luxembourg

Maria Gavrilidou
Institute for Language and Speech Processing, Athens, Greece

Dafydd Gibbon
University of Bielefeld, Faculty of Linguistics and
Literature, Bielefeld, Germany

Eva Hajicova
Charles University, Institute of Formal and Applied
Linguistics, Prague, Czech Republic

Ferenc Kiefer
Hungarian Academy of Sciences Research Institute for
Linguistics, Budapest, Hungary

Bente Maegaard
Center for Sprogteknologi, Copenhagen, Denmark

4. Post-Panel Discussion

People interested in the problems listed here, as well as in other similar issues, are encouraged to address their observations to the Panel Chair, also after the LREC conference. These remarks will be published at the WEB site given above.