A Multilingual Research Infrastructure

Bente MAEGAARDa1

a University of Copenhagen and CLARIN ERIC

Abstract. In this presentation we will focus on the scientific and technical goals for CLARIN, the contributions of the various communities and the strategies for the future.

Keywords. research infrastructure, multilinguality, CLARIN

1. CLARIN Objectives and Vision

The ultimate objective of CLARIN ERIC is to advance research in humanities and social sciences by giving researchers unified single sign-on access to a platform which integrates language-based resources and advanced tools at a European level. This shall be implemented by the construction and operation of a shared distributed infrastructure that aims at making language resources, technology and expertise available to the humanities and social sciences research communities at large.

The CLARIN vision is to create a sustainable infrastructure that will provide the research community with easy and lasting access to existing and future language resources and state-of-the-art tools, wherever they are located, thus enabling world class, innovative HSS research capable of crossing national, linguistic and discipline boundaries.

Language is at the heart of many disciplines in the Humanities and Social Sciences, be it as an object of study, an instrument for communication and expression, or a means to record information and knowledge. Thus language resources and the tools that can extract knowledge from them are at the basis of scientific discovery in many disciplines. The multilingual nature of Europe constitutes a special challenge here, but also a special opportunity as over the last decades Europe has become a world leader in multilingual language processing, and CLARIN will be building on that and will be sharing the knowledge accumulated.

The presentation will describe the various instruments which will be used for obtaining the goals, it will describe some of the challenges, and it will give some examples of existing services which can already be used.

1 Bente Maegaard, University of Copenhagen, Njalsgade 140, Copenhagen S, 2300 Denmark; e-mail: bmaegaard@hum.ku.dk