Preface

The Sixth Workshop on Treebanks and Linguistic Theories is being held in Bergen (Norway) on December 7–8, 2007. Previous workshops were held at Sozopol (Bulgaria), Växjö (Sweden), Tübingen (Germany), Barcelona (Spain) and Prague (Czech Republic). Although these workshops have been traditionally held in Europe, they draw global participation.

Treebanks are language resources that include annotations at the levels of linguistic structure beyond the word level. They typically provide syntactic constituent or dependency structures for sentences and sometimes functional and predicate-argument structures. Treebanks have become crucially important for the development of data-driven approaches to natural language processing and to linguistic research in general. There are a number of ongoing projects aiming at compiling representative and reusable treebanks for specific languages. In addition, there are projects that develop tools or explore annotation beyond syntactic structure and beyond a single language.

Experiences in building syntactically processed corpora show that there is an important relation between formal linguistic theory and syntactic annotation. Therefore the connection between treebank development and linguistic theories and paradigms merits attention. The workshop aims at being a forum for all researchers and advanced students working in the area of treebanking.

Erhard Hinrichs and Julia Hockenmaier accepted the invitation to be keynote speakers at the workshop. There were twenty-four submissions. Fifteen papers were accepted for presentation and are included in this volume. Six presentations were accepted for the poster session; their abstracts are also included here. The review procedure was anonymous and the decisions were taken on the basis of at least three reviews by members of the program committee; the final decisions were the responsibility of the program committee chairs: Koenraad De Smedt, Jan Hajič and Sandra Kübler.

Koenraad De Smedt