Preface

The Ninth International Workshop on Treebanks and Linguistic Theories (TLT9) was held at the University of Tartu in Tartu, Estonia on 3-4 December 2010 (see `http://math.ut.ee/tlt9/`). This marked the first time that it was held in one of the Baltic states. Dates and locations of the previous workshops are provided in a separate section.

Since its first edition in 2002, TLT has served as an ideal venue for new and on-going high-quality work related to syntactically-annotated corpora, i.e., treebanks. Treebanks have become crucially important for the development of data-driven approaches to natural language processing, human language technologies, grammar extraction and linguistic research in general. Additionally, there are projects that explore annotation beyond syntactic structure (including, for instance, semantic, pragmatic and rhetorical annotation), beyond a single language (for instance, parallel treebanks), and beyond simply written data, incorporating properties of speech. The papers for TLT over the years since 2002 have done much to capture this range of work, encompassing descriptive, theoretical, formal and computational aspects of treebanks.

Experiences in building syntactically-processed corpora have shown that there is a relation between linguistic theory and the practice of syntactic annotation. Since the practices of building syntactically-annotated corpora have illustrated that aiming at a more detailed description of the data becomes more and more theory-dependent, the connections between treebank development and linguistic theories need to be tightly connected in order to ensure the necessary flow of information between them.

The call for papers for TLT9 requested unpublished, completed work. 35 submissions were received, authored by 95 different authors, illustrating the extremely collaborative nature of these works. The submissions were authored by researchers from 16 different countries in America, Asia and Europe, and each submission was evaluated by three reviewers.

The Programme Committee consisted of 19 members (including the 3 co-chairs) from 10 different countries, all working as reviewers. Based on their scores and the comments they provided on the content and quality of the papers, 15 papers and 6 posters were accepted for presentation and publication. This corresponds to an acceptance rate of about 60%. The accepted submissions cover a wide range of topics related to both long-standing and new treebanks, reporting on aspects of their construction, querying, exploitation and evaluation.

Completing the programme are the invited lectures by Anke Lüdeling on "Syntactic Misuse, Overuse and Underuse: A Study of a Parsed Learner Corpus and its Target Hypothesis" and by Joakim Nivre on "Harvest Time – Explorations of the Swedish Treebank". Both of these invited talks connect years of treebanking with exciting new developments in learner corpora and the process of building larger treebanks.

Following in the tradition of TLT's recent editions, a co-located event was also organised (see `http://math.ut.ee/tlt9/aepc/`). This one-day event, preceding TLT9, was a Workshop on the Annotation and Exploitation of Parallel Corpora (AEPC). This co-located event arose from the consideration that parallel corpora are extremely useful data sets for which there is an increasing need of adding linguistic annotation. Syntax-enhanced approaches to machine translation rely on a strong connection to treebanks, and thus participants in both workshops benefitted from the unique combination of the workshops.

We wish to express our gratitude to the members of the programme committee, who worked hard to review all submissions under hard time constraints. We also want to thank University of Tartu, Institute of Computer Science and Estonian Center of Excellence in Computer Science for sponsorship and support to the workshop.

The TLT9 Co-Chairs

Markus Dickinson; Indiana University; Bloomington, IN, USA

Kaili Müürisep; University of Tartu; Tartu, Estonia

Marco Passarotti; Catholic University of the Sacred Heart; Milan, Italy