

# Invited talks

## **MOLTO: multilingual on-line translation**

Aarne Ranta, University of Gothenburg, Sweden

The World Wide Web is a globally accessible, multilingual source of information, which has created an urgent need for automatic translation. Some of this need is satisfied by tools such as Google translate, Bing translator, and Babelfish/Systran, which provide quick translations of any web page between a large number of languages (currently 57 in Google translate). While these systems are impressive and useful, they have the character of *consumer* tools, rather than *producer* tools. A consumer, who uses one of these systems to translate a document, does it at her own risk. But if the producer herself publishes a translation, she is responsible for its correctness. For instance, an international e-commerce company publishing offers and product descriptions would take too high a risk by translating them by any of the consumer tools. There is little hope in machine translation to achieve at the same time large coverage and high quality. In the resulting trade-off, consumer tools must opt for coverage, whereas producer tools should opt for quality. Fortunately, a producer of information is often in control of the content to be translated, so that unlimited coverage is not necessary. The MOLTO project (<http://www.molto-project.eu>) targets this need, and aims to show the feasibility of production quality translation for limited domains but a high number of simultaneous languages (up to 15 within the project's case studies).

The technology used by MOLTO is based on GF (Grammatical Framework), which is a programming language for multilingual grammars. The talk will explain how production-quality translation systems can be built rapidly and economically by using GF. The key concept of GF is a *multilingual grammar*, a grammar that uses an *abstract syntax* as an interlingua, a shared semantic structure for multiple languages. The languages are related to the abstract syntax by reversible mappings. To help writing these mappings, GF provides a Resource Grammar Library, which implements the morphology and basic syntax of currently 18 languages. While the translation systems built in MOLTO are essentially grammar-based, they will integrate statistical machine translation methods to improve coverage and to learn parts of grammars from data. GF and related tools are open-source software available for all major platforms and constantly developed by an international community.