

GF - Grammatical Framework

The core of a MOLTO translation system is a multilingual **GF grammar** where meaning-preserving translations are obtained as composition of parsing and generation via the **abstract syntax**, an interlingua. GF is a framework for interlinguas in which the basic linguistic details of languages, inflectional morphology and syntactic combination functions, are provided via the Resource Grammar Library.

MOLTO will further improve grammar engineering in GF by:

- ♦ **Integrated Development Environment** to use the RGL and to manage large projects;
- ♦ **Example-based grammar writing** support to bootstrap a grammar from a set of example translations.

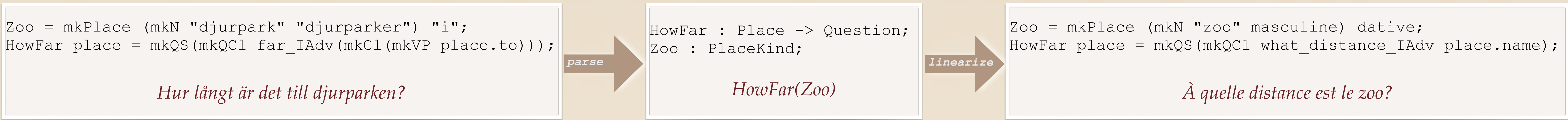
Statistical Machine Translation

MOLTO will develop and evaluate combination approaches to integrate grammar-based and SMT models in a hybrid, robust MT system. At least four variants will be studied:

- ♦ **baseline:** cascade of independent MT systems;
- ♦ **hard integration:** GF partial output is fixed in a regular SMT decoding;
- ♦ **soft integration I:** GF partial output, as phrase pairs, is integrated as a discriminative probability feature model in a phrase-based SMT system;
- ♦ **soft integration II:** GF partial output, as tree fragment pairs, is integrated as a discriminative probability model in a syntax-based SMT system.

OWL Ontologies

MOLTO sees ontologies as a way to formalize interlinguas in specific domains. Based on this observation, it will carry out research to develop two-way grammar-ontology interoperability that will bridge natural language and formal knowledge. The resulting MOLTO infrastructure will allow knowledge modeling, semantic indexing and retrieval using natural language. The engine will perform semi-automatic creation of abstract grammars from ontologies; derive ontologies from grammars, and retrieve instance level knowledge from/ in natural language by first transforming queries to semantic queries, and secondly by expressing the resulting knowledge in natural language.



MOLTO's goal is to develop tools for web content providers to translate texts between multiple languages in real time with high quality. Languages are separate modules in the tool and can be varied; prototypes covering a majority of the EU's 23 official languages will be built.



Multilingual On-line Translation

molto-project.eu
twitter.com/moltoproject



FP7- 247914 at a glance

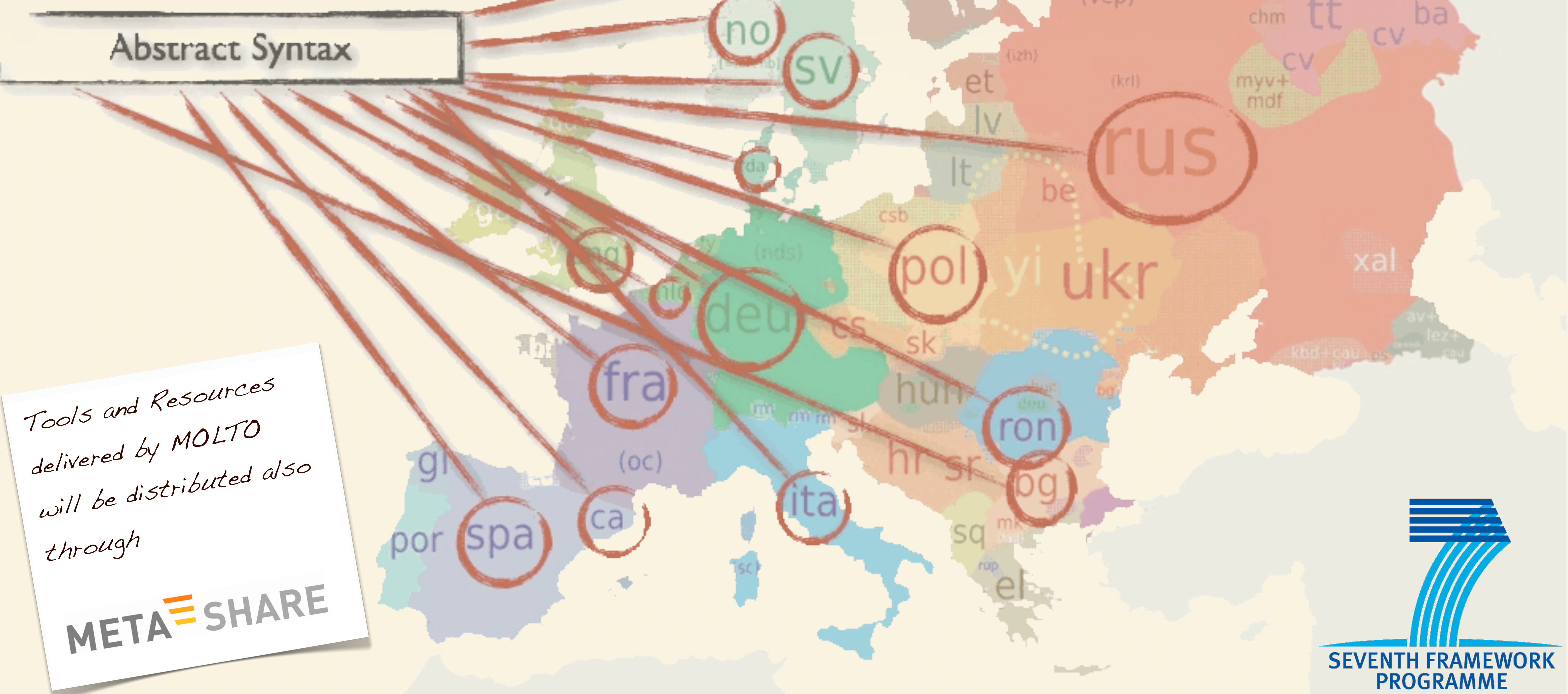
ICT-2009.2.2 - Language based interaction

Duration: **36 months**

Start date: **1 March 2010**

End date: **28 February 2013**

EU funding: **2.3 Millions €**



Tools and Resources delivered by MOLTO will be distributed also through **META SHARE**

