

WP10: Dissemination and Exploitation

Olga Caprotti
University of Gothenburg

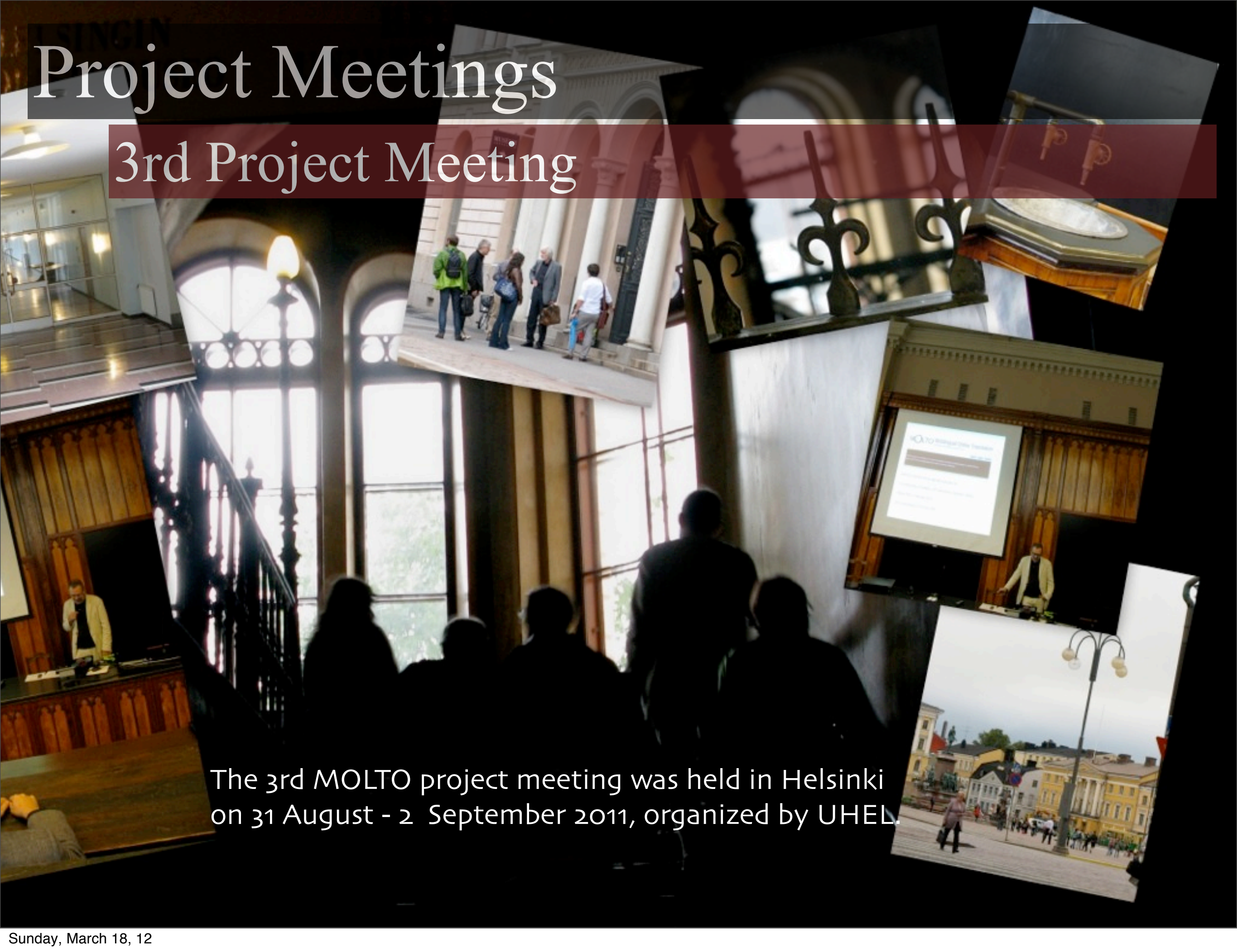
Jeroen van Grondelle
Be Informed

20 March 2012, UPC Barcelona
MOLTO Second Year Review

MOLTO

Project Meetings

3rd Project Meeting



The 3rd MOLTO project meeting was held in Helsinki on 31 August - 2 September 2011, organized by UHEL.

MOLTO Open Day

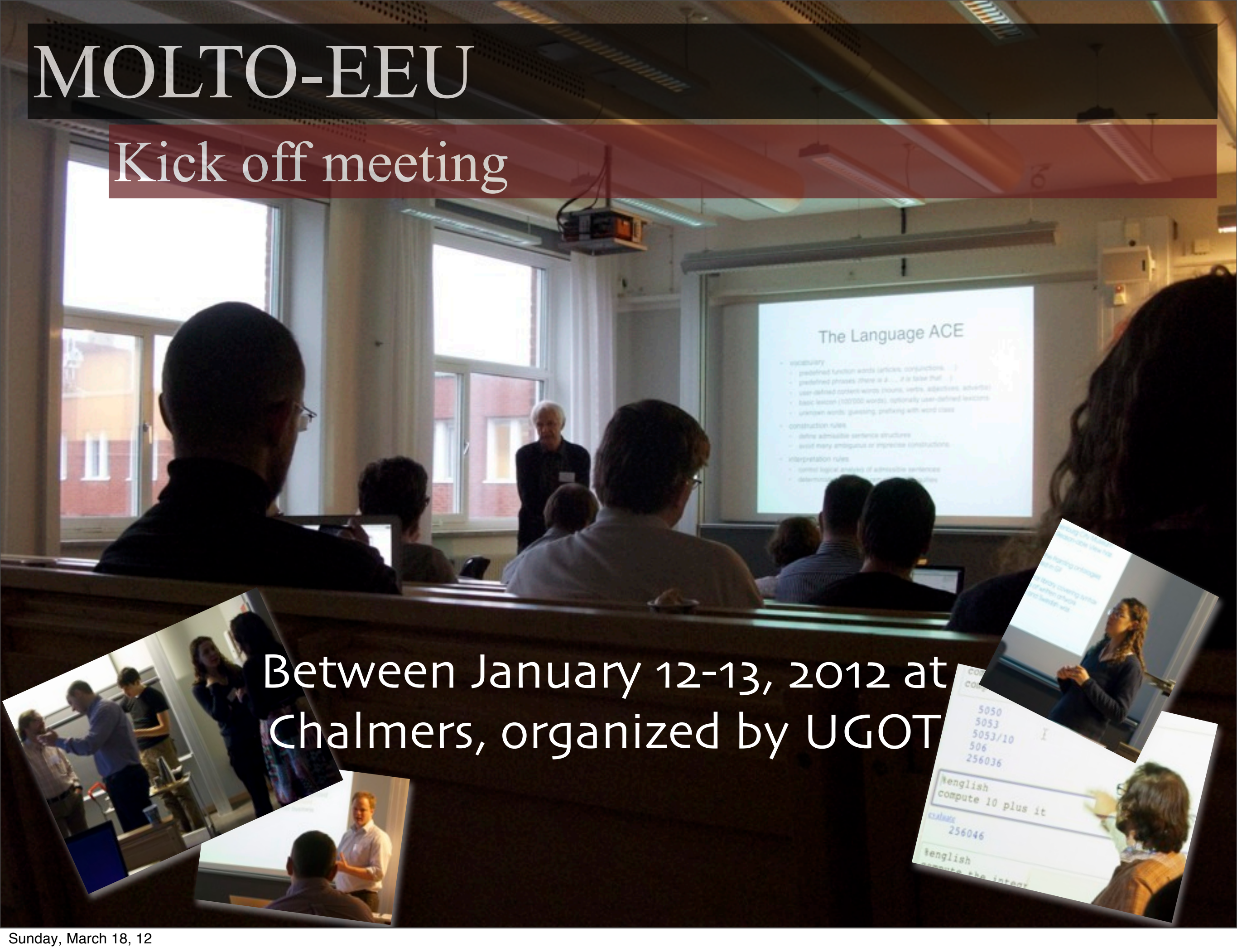
2 September 2012

Program

timeslot	title	speaker
9:00 - 9:30	[MOLTO Project: Overview and Half-Way Results](/node/1333)	Aarne
09:30 - 09:45	The Mathematical Grammar Library	Jordi
09:45 - 10:00	GF web IDE with focus on Finnish (demo)	Thomas/Aarne
10:00 - 10:30	Invited talk: Helsinki Finite-State Technology	Krister Lindén
10:30 - 11:00	Coffee break	
11:00 - 11:15	Term extraction from ontology with pre-constructed queries	Inari
11:15 - 12:00	TermFactory (demo)	Lauri C.
12:00 - 12:30	Machine translation evaluation	Maarit
12:30 - 13:00	Approximating term similarities with syntactic evidence	Seppo
13:00 - 14:30	Lunch at Café Arppeanum	
14:30 - 15:00	Invited talk: HFST as an open source environment for commercial and academic language tools	Kimmo Koskenniemi
15:00 - 15:30	A Reason-able view of Linked Data for Cultural Heritage	Milen Chechev (Ontotext)
15:30 - 16:00	Native PGF	Lauri A.
16:00	Closing	

MOLTO-EEU

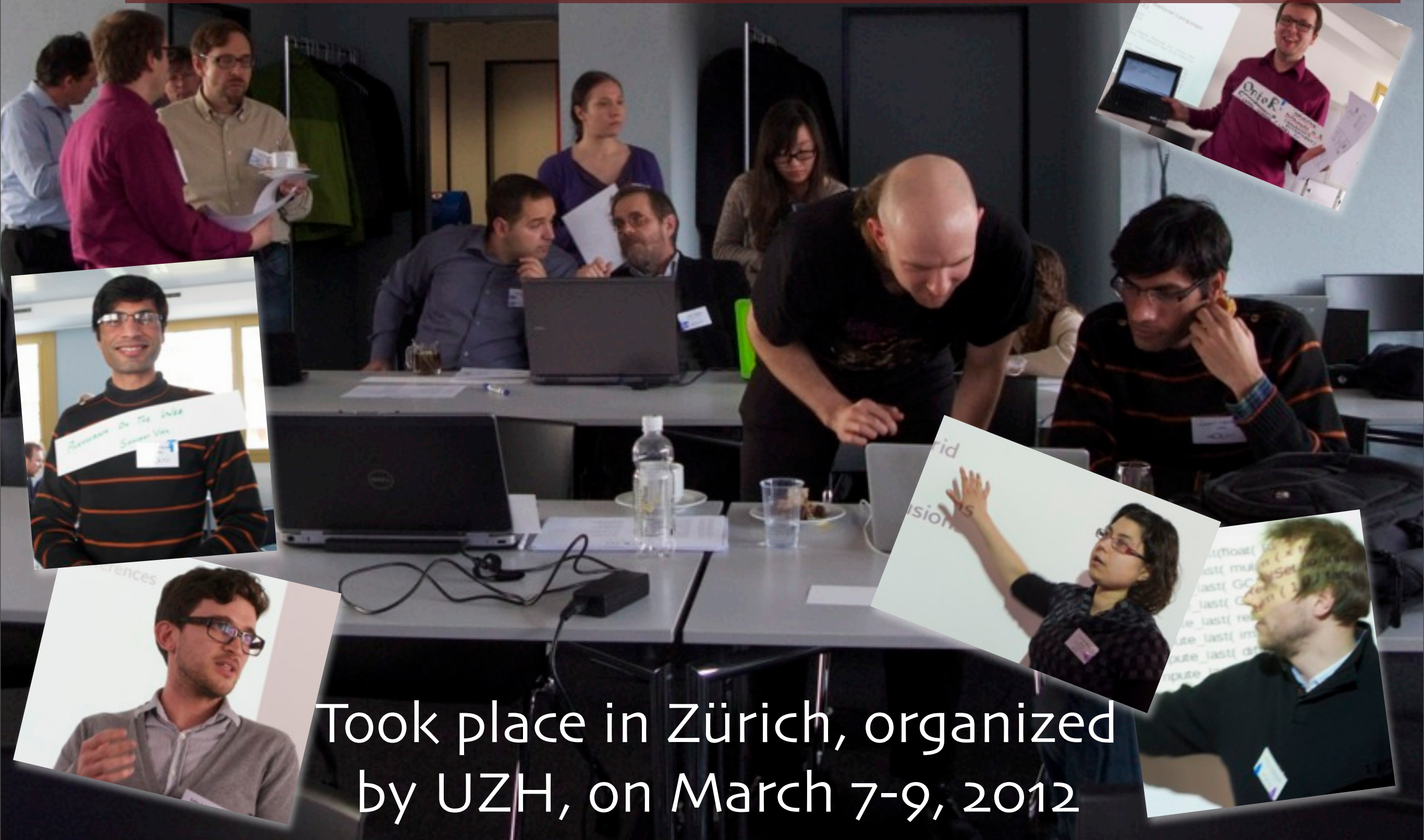
Kick off meeting



Between January 12-13, 2012 at
Chalmers, organized by UGOT

Project Meetings

4th Project Meeting



Took place in Zürich, organized
by UZH, on March 7-9, 2012

Publications

Conference papers, presentations, posters

Book

Grammatical Framework: Programming with Multilingual Grammars, Ranta, Aarne , CSLI Studies in Computational Linguistics, Stanford, p.350, (2011) [Google Scholar](#) [BibTex](#)

Conference Paper

Asiya: An Open Toolkit for Automatic Machine Translation (Meta-)Evaluation., Giménez, Jesús, and Màrquez Lluís , Fifth Machine Translation Marathon, Volume 94, Le Mans, (2010) [Google Scholar](#) [BibTex](#) Download: [PBML-2010-Gimenez.pdf](#) (170.97 KB)

Controlled Language for Everyday Use: the MOLTO Phrasebook, Ranta, Aarne, Enache Ramona, and Détrez Grégoire , Controlled Natural Languages Workshop (CNL 2010), Marettimo, Italy, (2011) [Google Scholar](#) [BibTex](#) Download: [everyday.pdf](#) (647.5 KB)

Creating Linguistic Resources with GF, Ranta, Aarne , LREC 2010, 05/2010, Valletta, Malta, (2010) [Abstract](#) [Google Scholar](#) [BibTex](#)

Document-level Automatic MT Evaluation based on Discourse Representations, Comelles, Elisabet, Giménez Jesús, Màrquez Lluís, Castellón Irene, and Arranz Victoria , 5th Workshop on Statistical Machine Translation (IWMT, 2010), (2010) [Google Scholar](#) [BibTex](#) Download: [wmt10-cgmac.pdf](#) (67.6 KB)

A Framework for Improved Access to Museum Databases in the Semantic Web, Dannélls, Dana, Damova Mariana, Enache Ramona, and Chechev Milen , RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING, 09/2011, Hissar, Bulgaria, (2011) [Abstract](#) [Google Scholar](#) [BibTex](#)
Download: [ranlpWS-final.pdf](#) (421.76 KB)

The GF Mathematics Library, Saludes, Jordi, and Xambó Sebastian , Proceedings First Workshop on CTP Components for Educational Software (THedu'11), 02/2012, Volume Electronic Proceedings in Theoretical Computer Science , Number 79, Wrocław, Poland, p.102-110, (2011) [Abstract](#)
[Google Scholar](#) [BibTex](#)

Publications

Conference papers, presentations, posters

Hybrid Machine Translation Guided by a Rule-Based System, España-Bonet, Cristina, Labaka Gorka, Díaz De Ilarraza Arantza, Màrquez Lluís, and Sarasola Kepa , Machine Translation Summit, 09/2011, Xiamen, China, p.554-561, (2011) [Google Scholar](#) [BibTex](#)

Download: [SMatxinT.pdf](#) (398.46 KB)

The MOLTO Phrasebook, Angelov, Krasimir A., Caprotti Olga, Enache Ramona, Hallgren Thomas, and Ranta Aarne , Third Swedish Language Technology Conference (SLTC-2010), 10//2010, Linköping, (2010) [Abstract](#) [Google Scholar](#) [BibTex](#) Download: [sltc2010.pdf](#) (145.36 KB); [MOLTO posterSLTC.pdf](#) (387.85 KB); [MOLTO posterSLTC.pages](#) (1.02 MB)

Multilingual Online Generation from Semantic Web Ontologies, Dannélls, Dana, Enache Ramona, Damova Mariana, and Chechev Milen , www2012, Lyon, France, (2012) [Abstract](#) [Google Scholar](#) [BibTex](#) Download: [www2012_submission_16-final.pdf](#) (1.22 MB)

Multilingual Packages of Controlled Languages: An Introduction to GF, Ranta, Aarne , CNL-2010 (2nd Workshop on Controlled Natural Language), 14 September, Marettimo, Sicily, (2010) [Google Scholar](#) [BibTex](#) Download: [cnl-2010.pdf](#) (1.11 MB)

Patent translation within the MOLTO project, España-Bonet, Cristina, Enache Ramona, Slaski Adam, Ranta Aarne, Màrquez Lluís, and González Meritxell , Workshop on Patent Translation, MT Summit XIII, 09/2011, p.70-78, (2011) [Google Scholar](#) [BibTex](#)

Download: [patentsMOLTO4.pdf](#) (282.42 KB)

The Patents Retrieval Prototype in the MOLTO project, Chechev, Milen, González Meritxell, Màrquez Lluís, and España-Bonet Cristina , WWW2012 Conference, Lyon, France, (2012) [Google Scholar](#) [BibTex](#) Download: [patentsretrieval.pdf](#) (397.01 KB)

Reason-able View of Linked Data for Cultural Heritage, Damova, Mariana, and Dannélls Dana , The Third International Conference on SOFTWARE, SERVICES & SEMANTIC TECHNOLOGIES (S3T), 09/2011, Bourgas, Bulgaria, (2011) [Abstract](#) [Google Scholar](#) [BibTex](#)

Download: [s3t2011_submission.pdf](#) (1.18 MB)

Robust Estimation of Feature Weights in Statistical Machine Translation, España-Bonet, Cristina, and Màrquez Lluís , 14th Annual Conference of the European Association for Machine Translation, EAMT-2010, (In Press) [Google Scholar](#) [BibTex](#) Download: [EAMT2010-EM.pdf](#) (188.23 KB)

Smart Paradigms and the Predictability and Complexity of Inflectional Morphology, Détrez, Grégoire, and Ranta Aarne , EACL (European Association for Computational Linguistics), 04/2012, Avignon, (2012) [Google Scholar](#) [BibTex](#)

Publications

Conference papers, presentations, posters

Tools for Multilingual Grammar-Based Translation on the Web, Ranta, Aarne, Angelov Krasimir A., and Hallgren Thomas , ACL 2010, July 2010, Uppsala, Sweden, (Submitted) [Google Scholar](#) [BibTex](#) Download: [molto-demo.pdf](#) (322.55 KB); [MOLTO posterACL.pdf](#) (982.89 KB)

Typeful Ontologies with Direct Multilingual Verbalization, Angelov, Krasimir A., and Enache Ramona , Controlled Natural Languages Workshop (CNL 2010), Marettimo, Italy, (2011) [Google Scholar](#) [BibTex](#) Download: [FinalSUMOCNL.pdf](#) (252.38 KB)

Using GF in multimodal assistants for mathematics, Archambault, Dominique, Caprotti Olga, Ranta Aarne, and Jordi Saludes , 02/2012, Digitization and E-Inclusion in Mathematics and Science 2012, (2012) [Abstract](#) [Google Scholar](#) [BibTex](#)
Download: [ArchambaultCaprottiRantaSaludes.pdf](#) (179.15 KB)

Events

GF Summer School



Frontiers of Multilingual Technologies

August 15-26, 2011

Barcelona, Spain.

Home

▼ About the school

Lecturers

News

Photos

Study materials

Calendar

▼ Programme synopsis

Getting started with GF

GF application
development

Panel discussion
recording

Presentations by
participants

Resource grammar
development

Theory

▼ Registration

Travel grants

Venue

Sitemap

The GF Summer School is sponsored by [CLT](#), the Centre for Language Technology of the University of Gothenburg, and by [UPC](#), Universitat Politècnica de Catalunya.

Forthcoming

FreeRBMT and CNL



[Call for Papers](#) | [About the workshop](#) | [Important dates](#) | [Participate](#) | [Programme](#) | [Proceedings](#) [På svenska](#)

Free/Open-source Rule-based Machine Translation Gothenburg, Sweden, 13-15 June 2012

This workshop aims to bring together the experience of researchers and developers in the field of rule-based machine translation who have decided to get on board the free/open-source train and are effectively contributing to creating a commons of explicit knowledge: machine translation rules and dictionaries, and machine translation systems whose behaviour is transparent and clearly traceable through their explicit logic. Also hybrid systems combining rule-based and statistical methods are in scope.

Programme Committee chair: [Aarne Ranta](#)
Organizing Committee Chair: [Olga Caprotti](#)

Proceedings to appear in
[CSLI Studies in Computational Linguistics ONLINE](#)



Third Workshop on Controlled Natural Language (CNL 2012)

29-31 August 2012
Zurich, Switzerland

A controlled natural language (CNL) is based on natural language but comes with restrictions on vocabulary, grammar, and/or semantics. The general goal is to reduce or eliminate ambiguity and complexity.

Some of these languages are designed to improve communication among humans, especially for non-native speakers of the respective natural language. In other cases, the restrictions on the language are supposed to make it easier for computers to analyze such texts in order to improve computer-aided, semi-automatic, or automatic translations into other languages. A third group of CNL has the goal to enable reliable automated reasoning on seemingly natural texts. Such languages have a direct mapping to some sort of formal logic and should improve the accessibility of formal knowledge representations or specifications for people unfamiliar with formal notations.

All these types of CNL are covered by this workshop.

Exploitation by Be Informed

WP12 is about adoption

Proof of the Exploitation Pudding

If succesful, Be Informed's verbalization component can be permanently moved to GF

Successful =

Quality texts in multiple languages

Modular integration with (multilingual) ontologies

Analysts are able to develop grammars

Future of GF (community) is secured

Adoption:

Be Informed will base its default verbalization on GF

Customers/Partners will develop custom grammars

Exploitation by Be Informed

Pre-requisites for industrial exploitation

Modular integration

GF can be adopted as an add-on, no low level conformance or changes required to adopt GF

Analysts are able to develop grammars

Grammar development discipline should align with one or more industry roles, preferably not linguistically trained

Future of GF (community) is secured

In many adoption models, GF is treated as a toolset, changing/fixing it is beyond typical industry expertise

Therefore, GF's future development and maintenance must be well grounded in a community to be widely adopted

Exploitation by Be Informed

Ideas for other forms of exploitation

Adoption of the Software Components

But also

- Verbalization service

 - Offering Newsfeeds from structured data

 - Objects, specifically (Social)Events

- Translation services

 - Like phrasebook app, in other domains

- Lexical services

 - Leverage all the knowledge in resource grammars

- Developing/selling grammars

 - Potentially based on standard vocabularies etc.