Interinstitutional Language Technology Watch: Language Technology at the European Institutions

Conrad Toft
IILTW Chair
Head of Translation Strategy
European Economic and Social Committee

conrad.toft@eesc.europa.eu
+32 2 546 8531
Agenda

- Interinstitutional cooperation & context
- Working methods
- Activities in 2010 - 2011
  - Areas of interest
- Where next?
Interinstitutional Cooperation

Executive Committee on Translation

Coordination Committee

Tools
Projects
Networks

Executive Committee on Interpretation

Coordination Committee

Ad Hoc Task Forces

INTERINSTITUTIONAL COMMITTEE FOR TRANSLATION AND INTERPRETATION
Context

- Framework for interinstitutional activities to
  - structure cooperative actions and
  - facilitate the information flow between institutions,
  - enabling a more permanent collaboration
  - towards common objectives or
  - even the launching of common projects
  - to pool research activities into efficient working methods and techniques
Membership

- Those dealing with LTW matters in their respective institutions:

- The chair of the II-LTW is held for 2 years, appointed by ECT following a CCT proposal
Mandate of the II-LTW

■ Vision
  □ To be the trusted forum for information exchange on language technologies for the EU institutions and bodies co-operating in the ICTI

■ Mission statement
  □ To identify global developments in the area of language applications which would be relevant for the ICTI language services

■ Goals
  □ Pool together information
  □ Forecast future technologies
  □ Disseminate information
Areas of interest 2010-2011

- Quality assurance
- XML authoring tools for non-translators
- Machine translation
- Conversion and comparison from other formats
- Workflow tools
- Electronic dictionaries
- Teleworking and translation
- Proofing tools for spelling and grammar
- Terminology extraction
Quality assurance (QA)

- Internal consistency within a document
- Consistency between source and target versions
- Commission has investigated ErrorSpy and QA Distiller, further tests needed
- Current generation of CAT tools include plug-ins for QA
XML authoring tools for non-translators

- Use of commercial tools in the institutions is currently fairly limited
- Confined to specialist purposes such as the production of website content
- Web-based amendment authoring tools under investigation at Parliament and Committees
- More scope for exchange of information on internal IT developments involving production or use of XML
Machine translation

- Exchange of information on tests carried out at Commission and Parliament
- Commission aims to have MT@EC service available as soon as possible
- Further information on machine translation will form part of the following presentation
Conversion and comparison from other formats

- Focussed on PDF conversion
- ABBYY, Nuance PDF Converter, Nuance Omnipage tested
- Optical character recognition also of interest
  - Particularly for identifying text in embedded images
Workflow tools

- Survey of current workflow tools
- Commission’s “TraDesk”, first step in integration of ICT components in document workflow
- In-house, commercial and open source alternatives to be examined
- Possible area for future collaboration
Electronic dictionaries

- Survey to be carried out on range of electronic dictionaries used
  - Recommendations
  - User satisfaction
  - Methods of delivery
  - Managing dictionary resources
- Possible area for further collaboration
Teleworking and translation

- Most already have teleworking
- Survey on
  - Current state of play
  - Technologies used
  - Limitations on teleworking
Proofing tools for spelling and grammar

- Linguistic tools included in Microsoft Office 2007 & 2010
  - spell-checkers
  - grammar and contextual spellers
  - new reference works
  - for writing assistance
Terminology extraction

- Exchange of information on testing of
  - Xerox TermFinder
  - Terminotix Synchroterm
  - SDL Multiterm Extract
  - Multcorpora Multitrans
- Questions of efficiency and inclusion in workflow
Where next?

- How to best serve new language combinations?
- How to manage multilingual source texts?
- How to maintain responsibility in the face of increased automated quality assurance?
- How to measure the impact of language technology on translation?
- Is there a role for external web-based machine translation technologies in EU institutions?
- What is the role of the translator in the new technological environment?
Summary

- Interinstitutional cooperation & context
- Activities in 2010 - 2011
  - Areas of interest
- Where next?
Any questions?