IN TERMINO QUALITAS

IN HUMAN AND MACHINE TRANSLATION

1st World Congress on Translation Studies
Workshop on Computer Assisted Translators vs Human Translation
Paris, Nanterre University, 11-12 April 2017
Overview

• The challenge of legislating in 24 languages
• The role of terminology for linguistic and legal consistency
• On-line cooperation between the EU institutions in terminology
• The role of terminology in CAT tools
• The role of terminology in machine translation and in quality assurance
The multilingual reality of the European Union
The European Union is the largest union of states, with a common body of legislation which is applied in numerous areas of life.

EU legislation is transposed into national law and applied in the official language of each country.

In most domains, a big number of the laws of the Member States are based on European legal acts.
The relationship between the European, national and local level leads to a mutual transfer of knowledge.

In the same way that legal systems, cultures and languages of the Member States have influenced the legal system of the Union, the latter redounds on the legal systems of the Member States and their linguistic and cultural environment.

The EU introduces terminology in the internal and external institutional network, which is then fixed by usage in the Member States and their customs, and fed back to the institutional level.

The fact that the official language of a Member State has at the same time become the official language of the Union has, in many countries, promoted the emergence of a conscious language policy.
Cultural diversity

‘EU legislation’ is based on a codecision system between representatives of the Member States: Commissioners, Ministers and Parliamentarians. At the Institutions, each of them reflects the culture of their own people.
The very first regulation of the European Union stipulated that all official languages, of which there were 4 at the time, were to be treated on a footing of equality and that all documents should be translated into those languages. This meant that each official language was recognised as original for each legal act.

Symbolically, the multilingualism regulation has retained its number – 1 – while the number of official languages has gradually increased, with the addition of further 20 languages.
The official languages and the working languages of the institutions of the Union shall be Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish and Swedish.
The EU and Multilingualism
Translation is present at every step in the legislative process; from the first draft, which originates with the Commission’s experts, through all the negotiations with national departments and with the Council and Parliament, to the ultimate legal act, on which a vote is taken in each national Parliament, and which then becomes an ‘original law’. 
In implementing the same legal acts in all States and in all the languages of the European Union, the most important point is that fundamental linguistic concepts should be understood in the same way everywhere.

This is only possible thanks to a shared and consistent terminology.

IATE is a concept-oriented database covering more than 100 fields.
Different uses of Terminology

- **Sentimental analysis**: the terms used in a text regarding capital markets can have an influence on the behavior of investors. Example: the translation of "hedge funds" in several languages using or choosing not to use the word “risk” as part of the term.

- **Terminology in journalism and communication**: the language used to communicate news and to pass messages is different than administrative language.

- **Adaptability of terminology according to the audience**: for some target groups a different level of language – example: medical terminology database of the "medecins sans frontières".
Interinstitutional communication and cooperation
Collaborative platform

Terminology Projects

Interinstitutional Cooperation

Term Projects by Institution
Bienvenidos a la wiki ES Term

Foro de terminología
- Terminos sobre migración, asilo y gestión de fronteras
- Migración, asilo y gestión de fronteras

Actas y resúmenes de reuniones

Registro de consultas

Guías y manuales

Recursos terminológicos

Lista de contactos

Calendario
External Communication
termcoord.eu
TermCoord APP
Ontology and semantic web
Quality assurance in CAT tools and MT
Normative terminology for proactive use
Terminology resources for post-editing
Terminology and artificial intelligence – bioethics
Technical aspects – Term extraction etc.
Interlinking of terminology – Cloud technology and Metasearch
### Added value of Terminology

<table>
<thead>
<tr>
<th>TRANSLATION MEMORIES</th>
<th>TERMINOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>How something was translated</td>
<td>How something should be translated</td>
</tr>
<tr>
<td>From the past</td>
<td>For the future</td>
</tr>
<tr>
<td>Variety</td>
<td>Consistency</td>
</tr>
<tr>
<td>Segment level (phrase)</td>
<td>Subsegment level (term)</td>
</tr>
</tbody>
</table>

B. Having regard to Rules 115(5) and 110(2) of its Rules of Procedure,

- whereas, due to the economic crisis, the overall unemployment rate had by October 2012 risen to the unprecedented level of 20.7 %, with 25.91 million people looking for work;
- Whereas, November 2012

- whereas the youth unemployment rate has soared to 23.4 %, leaving 5.06 million young people unemployed, in part reflecting mismatches between skill supply and labour market demands though often regardless of the education levels of job seekers; whereas research shows that youth unemployment often leaves permanent scars, such as an increased risk of...
IATE in CAT tools

• Termbase retrieval from IATE: interinstitutional tool for retrieving relevant terminology from IATE in termbase format to be used in CAT tools

• Term recognition mechanism: automatically marks relevant terms in the translation environment on the basis of IATE contents

• Autosuggest: automatically proposes term suggestions when words are being typed

• Add a term: allows users to save search results in a termbase for new terminology, and share them with all translators using the same termbase
There is a new kind of document-specific termbases which can be exported through a recently developed interface of IATE and allows for the retrieval of IATE data relevant for any uploaded text. This new interface, called Term Recognition Module, is now available in production for all terminologists and translators in the EP.

A Task Force within the IATE Management Group identifies data sets in IATE which can be used as prescriptive terminology.

The EP is developing its own CAT tool (CAT4Trad) allowing for the treatment of multilingual documents where the retrieval of IATE data will also be foreseen.
IATE Public provides now the possibility to download TBX files filtered by languages and domains. Users can extract subsets of data as follows, using the extraction tool IATExtract:
Terminology in Machine Translation

Sources: “Machine translation” on topsy.one
Terminology Workflow

- Automatic term extraction (Source Language)
- Terms in all languages
- Insertion in CAT tools and/or machine translation
- Translation

Central Terminology Coordination and/or terminologists
Terminologists
Pre-translation
Translators
Quality Control in Post-Editing

- Grammar, syntax and spelling
- Correct and consistent terminology
- Style according to the rules and audience targeted
- Use of same references as for manual translation (TM, glossaries, termbases, etc.)
Terminology in Voice Recognition

Bilingual term bases extracted from IATE are used for the direct translation of recorded speeches.
The European Parliament is creating an own CAT tool to cover the needs of translating multilingual documents and specially amendments containing modifications on several languages. This CAT tool called CAT4Trad is using a repository with several resources and among them terminology from IATE.
Thank you for your attention!