# Terminology as an Integral Element in the Translation Process

Prof Dr Frieda Steurs

Lessius/KULeuven

Empowered by *Lessius*

# Some acronyms

- TMM = Terminology Management
- TWF = Terminology WorkFlow
- KOS = Knowledge Organisation System
- I18n = Internationalisation
- QA = Quality Assurance
- SL = Source Language
- TL = Target Language
- ISO = International Standards Organisation
- ROI = Return on Investment

Empowered by *Lessius*

# Language Industry

- Translation, interpreting and consultancy
- Audiovisual translation
- Software localisation
- Localisation of products in multilingual settings (a.o.website globalisation)
- Translation technology : CAT, MT, etc.
- Language training (language courses made to measure, e-learning)
- Multilingual conference support

# Revenu of the Language Industry (EU)

**Sector values of the Language Industry (in million €)**



- Translation, interpreting
  Software localisation
  Website globalisation

- Language teaching

- Audiovisual translation

- Language technology development

- Conference support
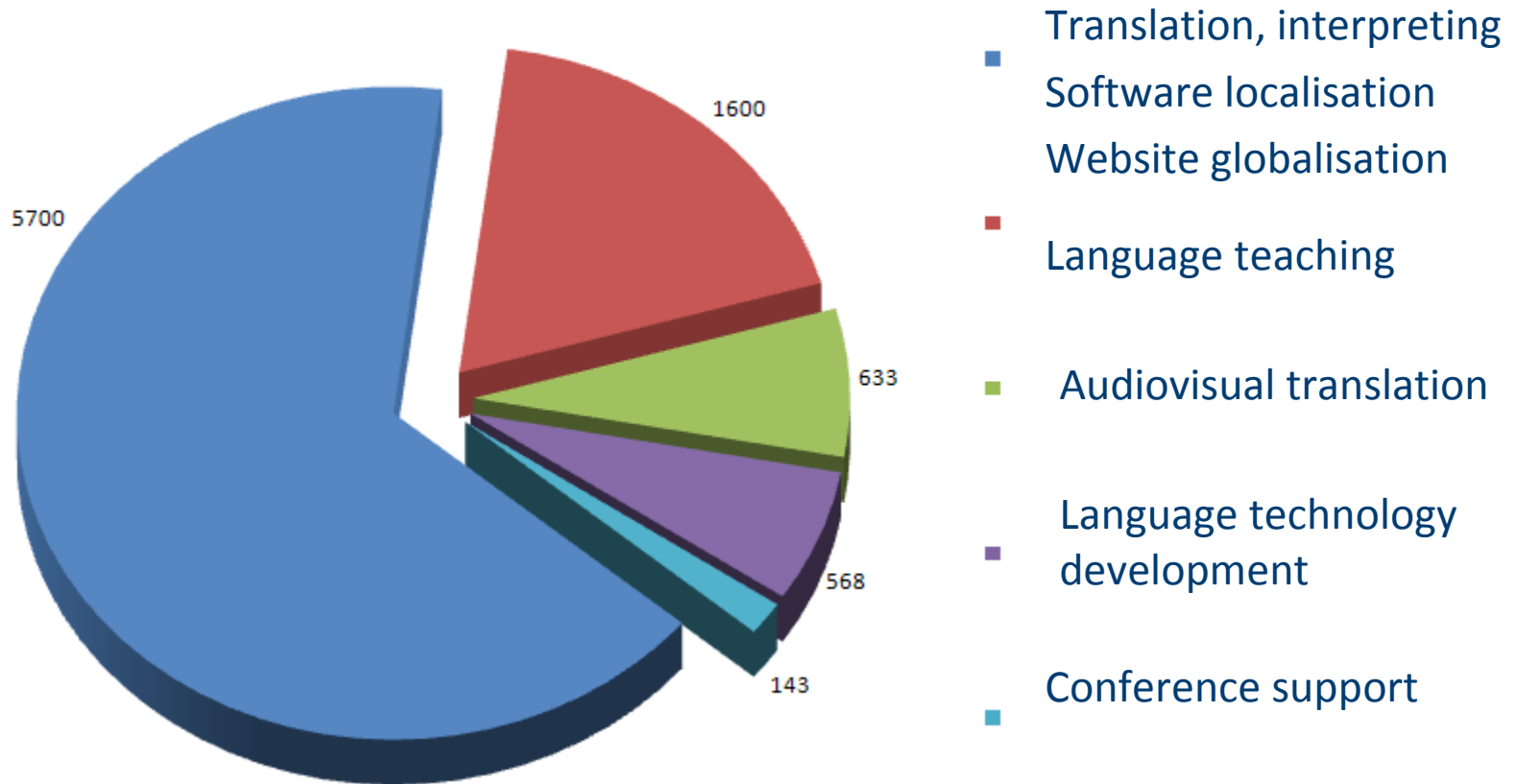
# The Language Industry (EU)

| Subsector (EU) | Revenue (€) |
|---|---|
| Translation, interpreting Software localisation Website globalisation | 5,7 billion |
| Audiovisual translation | 633 million |
| Language technology development | 568 million |
| Language teaching | 1.6 billion |
| Conference support | 143 million |

Empowered by *Lessius*

# Terminology as a core business

- Terminology is an important factor in a modern international business environment
- Knowledge driven society
- Knowledge is wealth, is the intrinsic asset of a company
- Knowledge is represented by language
- Very often : technical, scientific, legal, administrative language, driven by technical and complex concepts

- "There is no knowledge without terminology"

Empowered by *Lessius*

# The ongoing changes in vocabulary/terminology

- LSP (language for specific purposes) vs LGP (language for general purposes)
- Terminology vs general vocabulary
- Influence : mutual change and borrowing
- Constant evolution
- Shift in technicity of terminology, shift in use

# Word formation in LSP

- Productive splinters :
  - Watergate, Irangate, Pasquagate, ...
  - Brunch : breakfast and lunch
  - Hotel, motel, airtel,...
- Change in meaning / analogy
  - Terminal :
    - a) terminating point, part, or place
    - b) a station, usually at the end of a railway line
    - → computer systems
- Translation / equivalences / completely new terms

# English vs other languages

- English vs Dutch vs Afrikaans
  - Computer – computer – rekenaar
  - Milkshake – milkshake – melkschommel
- English vs Dutch (NL) vs Dutch (Flanders)
  - PINcode – pinnen – elektronisch betalen
  - Chip – chippen/chipknip – Proton
  - E-mail – e-mail – e-post (French : courriel)
- English vs French vs Dutch
  - Flat cable – cordon ombilicale – flat cable
  - E-mail – courriel/e-mail – e-post/e-mail

# IT

- *To retrieve* : ophalen, opsporen, terughalen, zoeken, terugzoeken, terugvinden

- *Benchmark* : norm, prestatietest, referentiepunt, standaard, benchmark

- *Bubble memory* : bellengeheugen, magnetisch geheugen, bubble geheugen, magneetbellengeheugen,

- *Workstation* : werkstation, terminal, PC ?

# Terminological confusion

- Unwanted, but unavoidable
- New terms are very often created in advertising and journals ; not through terminologists or linguists
- Confusion costs lots of money
- In-house terminology of companies : even a lot of inconsistencies there
- Awareness raising

F.

# The term

- Terms are specialised designations which designate clearly definable concrete or abstract objects
  - E.g. clutch, criminal proceedings, personal computer, LAN, etc.
- Object, concept, term : the triangle

term
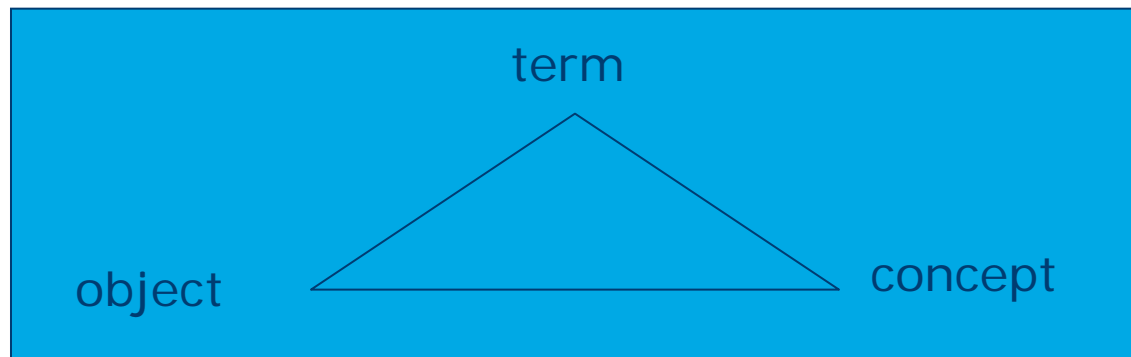
object                                    concept

# Concepts

- Comprise the more or less specific characteristics of particular individual objects (individual concepts : e.g. silver arrow) or a whole class of objects (general objects : e.g. racing cars)

- Definition and delimitation of the concept and determination of the concept in the system of concepts (semantic networks)

- Conceptual systems (shared, essential and deliminting characteristics)

# Terminography

- Registering the terms
- Creating terminology collections
- Specialised dictionaries
- Mostly done in an electronic way
  - Multifunctional
  - Updating of the datacollection
  - Levels of use

# Content management

Content management systems are systems that capture, archive, index, manage, author, combine, link, and distribute internal and external information to create a <span style="color:red">knowledge repository</span>

# The need for a centralised process

- Write it once
- Reuse it in many places
- Translate it once
- Reuse it in many places

# A whole spectrum

- Enterprise mission and statement
- Criticality of terminology for the core business
- Criticality of quality and branding concerns
- Recognition of tangible and intangible ROI
- Terminology user groups
- Buy-in by stakeholders (design, engineering, marketing, clients, top-level management)

- An important economic factor

Empowered by *Lessius*

# Enterprise types and stakeholders

- The nature of the enterprise
  - Government services
  - Industrial
  - Research institutes
  - Localisation/translation bureaus
    - Types of clients
    - Text type
  - Web content management providers
  - Freelance and in-house technical writers and translators
  - Librarians and knowledge organisation environments

Empowered by *Lessius*

# TMM : mission and strategic position

- Support for a national language policy
  - Catalan government, Flemish government, Gaelic…
- Support for global enterprise activity
  - Large companies have their own terminology lists (in-house terminology)
- Support for technical writing and translation
  - Also for free-lance translators
- Support for defining global concepts
  - In international organisations : internationalisation / localisation

- TMM approaches :
  - Prescriptive vs descriptive
  - Ad hoc vs systematic
  - Text–driven vs subject field driven

Empowered by *Lessius*

# TWF Task Issues

- Language planning
- Standardisation
- Document production
- Controlled language document production
- Activities in localisation and multilingual documentation environments
- Support for machine vs human-oriented translation
- Content management in a dynamically changing Web environment
- etc.

Empowered by *Lessius*

# Knowledge oriented TWF

- Enterprise related information and knowledge management
  - Terminology as a function of taxonomy, ontology and information retrieval (KOS : knowledge organisation systems)
    - Classifications in companies :
      - » E.g. *Hyundai  industrial equipment (wheel loader):*

        *Engine/Electrical system/Power train system/Hydraulic system*

        *e.g. transmission*

        *manual - automatic*

Empowered by **Lessius**

# Knowledge-oriented TWF

- KOS interacts with :
  - Monolingual technical writing and product development
    - e.g. when writing softwaretools : helpfiles, text in the menu, etc.
  - Translation and localisation
  - Multilingual technical writing
  - E-business solutions (taxonomy in websites, catalogues, e-sales, etc.)
  - Inventory control and logistics
  - General information retrieval and processing
    - Keywords, indexing systems

Empowered by *Lessius*

# Cost effectiveness

- How can you calculate the ROI?
- Terminology management activities cost time and money.
- It is very timeconsuming, and in the beginning, the results are not visible nor measurable
- It is an investment
- Is it feasible for individual translators?
- Is it feasible for groups working together?
- Is it feasible for bureaus and translation services in government and industry?
- Should it be limited to translation services, or should it penetrate deeper into the company or government services?

Empowered by *Lessius*

# Terminology : a critical factor

- Potential for market losses (no effective translation, problems with localisation)
- Potential for communicative losses : mistakes, not ready for possible markets, ...
- Liability claims, risk of product failure, human injury (Medtronic, car brands)
- Marketing issues (branding in foreign markets)
- Relative significance of terminology
  - To the process
  - To the product
  - Terminology is more critical when you are dealing with complex texts  (e.g. legal, administrative systems) or if you sell "intelligent products" such as software, computers, medical equipment etc.

Empowered by *Lessius*

# Medtronic

- http://www.medtronic.com/

- Pacemakers
- Neurostimulators
- Stents (heart)
- Etc.

- Critical material
- Terminology management /senior terminologist / multilingual management

Empowered by *Lessius*

# ROI

- Variable vs fixed costs
- TMM tasks are very much hidden in general overhead : many people participate in terminology management, without even noticing
- Cost of not agreeing to uniform terminology is hidden in overhead costs
- Cost of correcting mistakes or recouping damage to branding is difficult to calculate
- Compounded costs due to the persistence of defective communication

Empowered by *Lessius*

# Maximum efficiency in TMM

- Maximum efficiency in the composition of multilingual contents begins with the source text. For example, if the source text uses 200 different renderings for 100 technical terms – e.g., "hard drive" in one text and "hard disk" in another – translations into 10 languages would result in 2,000 target-language terms instead of the needed 1,000 terms.

- For one thing, this impairs the consistency and quality of the documents. Even worse, it generates substantial extra expenses that are multiplied by the number of languages.

Empowered by *Lessius*

# Source text control

- Therefore, the objective is to compose the source text using wording that complies with defined linguistic standards and for which translations are already available.

- For instance, if the sentence "Please observe the safety instructions" already exists in the translation memory, using this sentence enables the use of the corresponding translation from the translation memory; in contrast, using "The safety instructions must be observed" would call for a new translation

Empowered by *Lessius*

# Cour de cassation Belge

- Thèse
  - Opinion doctrinale ou position adoptée par un plaideur (thèse de la défence)

  - Stelling (12x)
  - Hypothese (1x)
  - Zienswijze (2x)
  - Houding (1x)
  - Stellingname (1x)
  - Standpunt (3x)

Empowered by *Lessius*

# Fixed and visible costs for TMM

- The costs of doing systematic terminology work
  - Are considerable
  - Are up front costs
  - Can be calculated using traditional means

- ROI
  - Is less visible
  - Calculations show that a term costs minimum 1000 Euro
  - Mistakes ? Hard to calculate
  - Spreading of inconsistencies and mistakes? Even harder to calculate

Empowered by *Lessius*

# Calculating ROI

- Greater general applicability of specific terminology units (greater frequency) = the greater the return on terminology management costs
  - E.g. banking sector (Brussels) :
    - Translation services create a terminology database for the purpose of translation and multilingual document management
    - The database is concept oriented, and can be used on all levels in the bank
    - The database is a very useful tool to train new employees in the typical in-house terminology of banking products in 4 languages.
    - This type of expansion and multiple usability makes the ROI considerably higher

Empowered by *Lessius*

# Calculating ROI

- The greater the quality, safety and competition related criticality = the greater the return
  - E.g. car constructors
    - Mistakes in the manuals : problems with liability claims, safety procedures
    - Usage of the terminology not only for owner manuals, but also for technical manuals , training courses for technicians, etc.

- The greater the degree of integration between CAT, TM and MT, the greater the payback of the effort
  - Terminology databases used as assets in all computer applications for writing, editing and translating documents

Empowered by *Lessius*

# Terminology Management Products

- Integrated and non-integrated workflow models
- Varying complexity with respect to the data model
- Varying use of terminology tools in
  - Content documentation
  - Concept definition (ideally by subject experts in the source and the target languages)
  - Terminology product delivery
    - On-screen lookup
    - Web based resources
    - Hard copy output

Empowered by *Lessius*

# Complexity of the data model

- How much data can you afford?
- A continuum of practice in a range defined by :
  - Creation of a non-systematic glossary or spreadsheet (text and job oriented)
    - SL term and TL term
    - POS, Gender, and/or a note
  - Rich complex entry input models
    - Subject field classification
    - Client designation
    - Concept – oriented definition
    - SL term/TL term/POS/ termtype/ register/context etc.
    - Documentation : citations, graphics, etc.

Frieda Steurs  - June 2010

Empowered by *Lessius*

# COTSOES : Conference of Translation Services of European States

- Recommendations for terminology work
- Terminology as an economic factor, a production and marketing factor
  - New requirements for commercial information and documentation
  - Product declaration and tendering in the language of the customer
  - Multilingual documentation and stock-keeping

Empowered by *Lessius*

# Cotsoes

- Terminology as an economic factor with regard to safety, quality and profitability
- Small and medium sized companies face considerable difficulities to manage this. Terminology work is relatively costly.
- It is absolutely necessary to maintain their position in national and international competition (terminological requirements imposed by reliable product information or documentation must be met)
- Efforts into harmonising terminology of a subject field and setting terminological standards.

# ROI

- ROI = $\dfrac{\text{revenues} - \text{costs}}{\text{investment}}$

  Revenue enhancement case

  Revenue protection case

  Cost reduction case

  Cost avoidance case

Empowered by *Lessius*

# ROI in terminological activities

- Litigation avoidance
- Fulfilling legal obligations
- Quality of patents
- More efficient communication
- Avoidance of after-sales issues

Empowered by *Lessius*

# ROI

- Possible losses :
  - Missed launch on a market for a particular Xmas action : documentation for a smartphones series was not ready ; localisation was delayed for 2 months : 17% loss of potential from local market
  - Local document delivery is necessary to make maximum profit on a market
  - Litigation problems : product recalles, damage compensations, ...

Empowered by *Lessius*

# Terminology in SL and TL

- Source language terminology management is a "best practice" to take out confusing language at the beginning of the process

- Target language terminology management is a "best practice" to prevent multiplication of confusion into target markets

- Target language terminology management is a means to preserve brand dillution in target markets across global enterprise content (localisation)

Empowered by *Lessius*

# Problems in many companies

- Global markets :
  - Very often high volumes in technical documentation
    - E.g. 17.000 source pages , 78.000 translated pages
    - Average of 50% re-use between products
    - All re-use is manual
    - Inconsistent use in terminology
    - Same concepts translated differently on the webpages, the brochures, the technical documents, the spare parts catalogues, etc.
    - Inconsistent structure in the technical documentation (lack of taxonomy, KOS)
    - No uniform processes and methods for authoring, translation and publishing

Empowered by *Lessius*

# Multilingual communication :expensive and ready for optimisation

- Also looming on the global content landscape is the ever-present mandate to reduce costs. This manifests itself in the need to streamline and eliminate redundancy in business processes, to automate manual tasks and to achieve visibility in project management. Translation, a highly manual process, is ripe for this type of process improvement. (The Gilbane Group, *A Market Report on Global Content Management 2006*, http://www.gilbane.com)

- The ongoing challenge is to strike the right balance between providing a satisfying customer experience in local languages/cultures vs. maintaining and enhancing one's brand. (The Gilbane Group, *Global Content Management: HP Talks the Talk of Worldwide Business, http://gilbane.com/case_studies.html*)

Empowered by *Lessius*

# What is workflow?

- Workflow is an operational aspect of a work procedure
  - How tasks are structured
  - What their relative order is
  - How they are syncronised
  - Who performs them and where, when
  - How information flows support these tasks
  - How these tasks are being tracked
  - Aspects of quality control can be added at all times

Empowered by *Lessius*

# TWF

- Terminology work flow can be a non-detailed item in a total workflow of document management

- TWF can be a particular task during the project planning

- TWF can be part of project oriented work or can be part of enterprise related global terminology management

Empowered by *Lessius*

# Resources

- Resources for TMM
- Raw source texts
  - Frequently flawed (esp. in localisation tasks)
  - Subject to ambiguity : polysemy, synonymy
  - Term extraction from the source text
    - Human or machine term recognition
    - Synonym identification and disambiguation
  - Determining target language equivalents

  - Resources :
  - Write once, read many times
  - Translate once, read many times
  - Central data repository

  - *"Where is the wisdom we have lost in information?"*

Empowered by *Lessius*

# Possible positions of TMM

- Very often : ad hoc TMM

- Reactive project specific TMM

- No influence on document production , no influence on SL or I18N

- Sometimes :

- TMM as a function of QA management (see the case of Yamagata)

- TMM and QA can become part of the planning stage of a project

- Proactive TMM (first corpus management, then terminology management, then other document management processes).

Empowered by *Lessius*

# Terminology management transactions

- ISO 12620 Computer applications in terminology -- Data categories
- Proposal of a term (extraction)
- Origination : creation of a term entry
- Data input
- Modification : update and correction
- Check by terminologist
- Verification by subject field experts
- Approval / withdrawal : chief terminologist
- Publication : master database, internet, intranet, hard copy
- Use in the field
- Modifications, suggestions
- Standardisation and language planning

Empowered by Lessius

# Tasks and specialists

- Project developers (engineers, designers, other subject specialists)
  - Terminological inconsistencies
    - e.g. the case of Ricoh Europe
  - Uncontrolled coinage of neologisms
- Terminology project group representatives
  - Technical writers
    - Frequently unaware of multilingual issues
    - Language specific problems (untranslatable items)
  - Translators
  - Knowledge engineers
  - Terminologists

Empowered by *Lessius*

# Who does what

- Master terminologist
  - Authority to change the master file
  - Authority for final approval
  - Concept oriented / source language oriented
- Language specific terminologist
  - Source and target language research
- Data input specialist (IT)
- Subject field specialist
  - Source and target language
- Expansion of the tasks according to the needs

Empowered by *Lessius*

# TMM solutions for technical writers

- Ontology enabled terminology delivery
- Style sheets
- Automatic term use checkers
- Issues :
  - Resistance to controlled language and controlled terminology
  - Coinage of translation-unfriendly terms
  - Coinage of "cute" but unmotivated terms

Empowered by *Lessius*

# TMM solutions for translation environments

- Terminology database integrated with translation memory and text production software

- Concordance features for access to terms in context

- Multiple layers of information

- Quick retrieval of basic information :
  - SL – TL + grammatical information

- Optional view of full terminological entries

Empowered by *Lessius*

# Workflow management systems

- SDL workFlow Service™
- Trados Global Content Management™
- Star Translation Workflow Server™
- Lionstream Workflow™
- LTC Organiser Multilingual Management and Workflow Control Software System™

Empowered by *Lessius*