Translation-oriented terminography in the electronic age

Dr Mariëtta Alberts
Pan South African Language Board
Outline of presentation:

- Introduction
- Terminology development
- Harvesting of terminology
  - Subject-oriented terminography
  - Translation-oriented terminography
  - Community-oriented terminography
- Electronic devices for term harvesting
- Conclusion
Introduction

- Information is distributed and knowledge acquired through terminology.
- The terminology of each subject field/domain is increasing with every new discovery, invention or development.
- The supplying of appropriate terms in various subject areas/domains should be a national priority – especially in a multilingual dispensation.
- Multilingual terminologies should be developed to equip South Africans with effective communication tools in all spheres of life.
Terminology development

- Terminology plays a pivotal role in language development and the promotion of multilingualism.
- Standardized terminology contributes to the quality of translation, editing, and interpreting services, in dictionary compilation and subject-related communication.
- Streamlined translation and interpreting services provide competitive advantages.
Information transfer, assimilation and retrieval should be through the first language (L1) or mother tongue.

It is a proven fact that information is best acquired (decoding process) and conveyed (encoding process) through the L1.

Standardized terminology in L1 leads to exact communication and misunderstanding or misinterpretations are avoided.
Harvesting of terminology

Subject-oriented terminography
- According to traditional subject-oriented principles and practice.
- Point of departure: conceptual domain
- SL terms are documented, defined, etc. and then TL equivalents are provided

Translation-oriented terminography (TOT)
- Point of departure: SL text to be translated

Community-oriented terminography
- Field work (e.g. indigenous knowledge systems; rural and urban speech communities, ...)

M Alberts, AFRILEX 2008, Stellenbosch
What is TOT?

- TOT standards had been developed and approved by the International Organization for Standardization (ISO) – ISO/TC 37 namely ISO 12616:2002 (Edition 1).
- This was also adopted as a South African National Standard: SANS 12616:2004 (Edition 1).
- Translation-orientated terminography (TOT) could be defined as the creation of topic-specific parallel corpora which result in bi- or multilingual terminological products containing all terms that occur in the source text (ST) and the translation equivalents in the target text (TT).
How does TOT work?

TOT is a methodology which will allow translators and terminologists to work together.

Parallel texts are used, i.e. source language (SL) and target language (TL) texts.

Various terminographical processes need to be taken into consideration, i.e. primary and secondary term creation phases.
Translation-oriented terminography

- SL and TL texts are **aligned**.
- SL and TL terms are **harvested**.
- These harvested SL and TL terms are **stored** in a glossary.
- The SL and TL terms form the **basis** for translation work between the relevant SL and TL.
- SL and TL texts are stored in a **translation memory** and can be utilized in future translations.
- **Result**: **standardized usage** of SL and TL terminology.
Vigorous plants produce masses of brightly coloured, large flowers over a long period. For beds, hanging baskets, containers and pots.

Annual

Groeikragtige plante lewer massas groot, kleurvolle blomme oor ‘n lang tydperk. Vir beddings, hangmandjies, houers en potte.

Eenjarig

Result: Glossaries

annual bed pl -s
colour n.
colourful container flower pl –s
hanging basket period pot pl -s
eenjarig bedding mv. -s
ekleur
kleurvol houer mv. -s
blom mv. –me
hangmandjie tydperk
pot mv. -te
TOT production of term lists:

- Study texts to align sentences
- Identify and match SL term with TL equivalent
- Extract key SL terms and TL equivalents
- Record SL and TL terminological information according to style guide (predetermined, standardized format)
- Supply definitions
- Discuss with established domain-specific terminology committee
- Finalize definitions
- Authenticate terminology (language-specific)
- Update database, publish, disseminate
Advantages of TOT:

- Continuous **recording** of terms and neologisms
- Quick and easy **retrieval** of reliable (verified) terminological information
- Constant **use** of available terminology
- **Standardization** of terms through frequent usage
- Continuous **support** and **interaction** between translators and terminologists
- Usage of any language as source language
- Streamlined terminology harvesting process (**data collection**)
- Facilitate terminology **coordination process**
Collaboration

- Terminologists and translators
- Language Research and Development Centres (LRDCs)
- National Lexicography Units (NLUs)
- National Language Bodies (NLBs)
  - Technical Committees: Terminology Development (verify and authenticate)
- Tertiary institutions
  - Subject specialists
  - Linguists
- Linguistic communities
Terminology and information flow in a developing community

IN

Data processing (i.e. language service)

OUT

End products to target users:
- dictionaries
- glossaries
- thesouri

Spelling and orthography rules

Term bank
Text bank (with headwords)
Libraries (document with headwords)

Information in:
- Curriculae
- Literature
- Textbooks
- Manuals
- Technical journals
- Annual reports, ...

Language practitioners (terminologists, lexicographers, translators, editors, interpreters, …)

(development of indigenous (TL) languages

Subject specialists, Administration, Education, Academics, Industry, Publishers, Mass media, Technical writers, Information experts, Standards organizations, …

- Excerpting of terms and words
- Preparing term- and wordlists
- Processing comments on lists
- Editing of products
- Printing of dictionaries/ lists

Register of translated texts
Register of terminologists, lexicographers, translators, editors, interpreters, …
TOT management:

- Software – Trados / TshwaneTerm / Wordfast / ...
- Database – national term bank or NLUs?
- Infrastructure
  - Translators/Terminologists/LRDCs/Language Units
- Verification and authentication
  - Subject specialists
  - Linguists
  - NLBs
- Dissemination
  - Publications (hard copy/ CD-ROM)
  - On-line
  - Intranet, Internet

M Alberts, AFRILEX 2008, Stellenbosch
Electronic devices and programs

- TRADOS
  - Translators’ Workbench
  - WinAlign
  - MultiTerm

- Wordfast
  - Translation memory
  - Compilation of glossaries

- TshwaneTerm/ TshwaneLex
  - Data capturing
  - Dissemination of information through hard copies, CD-ROM, Intranet, Internet

- Other software …
- Compatibility of software
- HLT endeavours
Conclusion

- Translation-oriented terminography requires dedication
  - Frequent and ongoing harvesting of terminology (and related information) in SL texts and TL translations
- Various collaborators will have to make inputs
  - Translators, terminologists, subject specialists, NLBs, linguistic communities (rural and urban), HLT
- TOT is a sure way of developing multilingual and polythematic terminology and the SA languages into functional languages.
Conclusion …

- Terminology is a **strategic resource** in terms of the functional development of languages.
- Effective subject-related **information transfer** and assimilation of knowledge occur through the use of **correct** and **standardized terminology**.
- **Electronic devices** and **programs** should be utilized to enhance the process of **terminology harvesting, documentation and dissemination**.
- Available and standardized terminology will bring about **exact communication**.
THANK YOU!

Dr Mariëtta Alberts
PanSALB
Private bag X08
ARCADIA, 0007

marietta@pansalb.org.za
Tel: 012 341 9638
Fax: 012 341 5938