

# **Machine Translation and Controlled English**

*Herman Caeyers, LANT*

## **Abstract**

In this lecture it will be shown how machine translation can be integrated successfully in the documentation flow by controlling the process already in the authoring phase. Simplified English is normally used in order to improve the readability of documents but it can also serve the translatability and more specifically the integration of machine translation. The experience gathered in the CASL project LANT is doing for General Motors in the US will be outlined.

Moreover, the use of machine translation technology in the development of a checker for controlled English, LanTmaster, will be discussed.

## **Dr. Herman Caeyers**

Dr. Herman Caeyers obtained a Master Degree in Theoretical Physics at the Katholieke Universiteit Leuven in 1978. In 1985 he promoted at the same university to Doctor in Mathematics with a study about mathematical models for mental processes which are the basis of music perception. He partly did this study at the Stanford University in California.

In 1985, Siemens Belgium asked him to lead the METAL project on Machine Translation. In 1990 he became the manager of the R&D department of Artificial Intelligence. Later on, he worked in Munich for Siemens-Nixdorf where he was responsible for the European projects in the domain of computational linguistics.

In 1992 Dr. Caeyers was Deputy Director of the EuroLang department of Sonovision Itep Technologies in France where he was advisor for the French Government in the domain of multilingual communication.

End 1993, he founded the company LANT in Belgium where he is the managing director.

## **LANT**

LANT was incorporated in 1993 by Herman Caeyers with the objective to become a significant player in the high-tech market of the development and the implementation of computer based Natural Language Processing products (NLP) ranging from electronic dictionaries to automatic machine translation (MT). The authorized capital of LANT has been raised recently to \$US 3.2 million.

LANT is a knowledge based company specialized in the development of solutions in the field of multilingual documentation and communication. In April 1996, LANT obtained the usage, sales and development rights as well as the final ownership of the total METAL technology, based on a contractual agreement with Siemens-Nixdorf. The METAL technology and its further development is marketed under the LANT@MARK™ trademark. LANT also took over all the rights of the Eurolang Optimizer translation memory from Sonovision Itep Technologies. LANT is also incorporated in the US.

LANT concentrates on two international markets: language related activities within large and middle-sized organisations and corporations as well as on the Internet world.

Herman Caeyers

LANT NV

Interleuvenlaan 21, 3001 Leuven

Belgium

Phone: +32 16 405140

Fax: +32 16 404961

E-mail: [Herman.Caeyers@lant.be](mailto:Herman.Caeyers@lant.be)

## Company Milestones



- ✦ **Founded end 1993**
- ✦ **March 1996: Lant obtained development and sales rights and the future ownership of the Metal Technology from Siemens-Nixdorf Germany**
- ✦ **April 1996: Lant obtained the development and sales rights of the French and Dutch Metal lingware components from Siemens-Nixdorf Belgium**
- ✦ **May 1996: Lant acquired the CASL project of General Motors Detroit**
- ✦ **April 1997: Lant obtained the exclusive rights of the Eurolang Technology from Sonovision Itep Technologies**

EAMT Copenhagen 97

## Company Profile

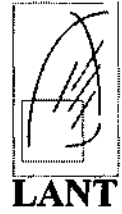


- ✦ **Market: multilingual documentation and communication**
- ✦ **Project oriented**
- ✦ **Further development of basic modules**
- ✦ **Integration of these modules into market solutions**
- ✦ **Customisation of the solutions**

**Based in Leuven (Belgium)**  
**Incorporated also in New Hampshire (United States)**

EAMT Copenhagen 97

# Product Strategy



## Documentation

## Communication

The Document Based  
Solution

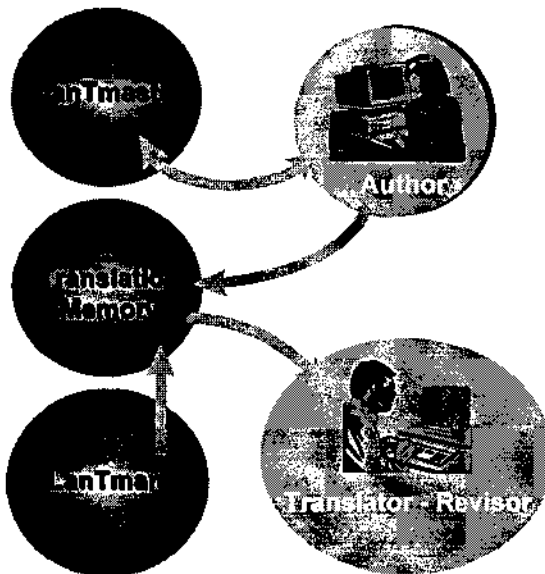
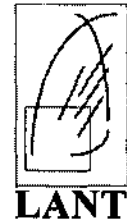
Internet

The Object Based  
Solution

Intranet

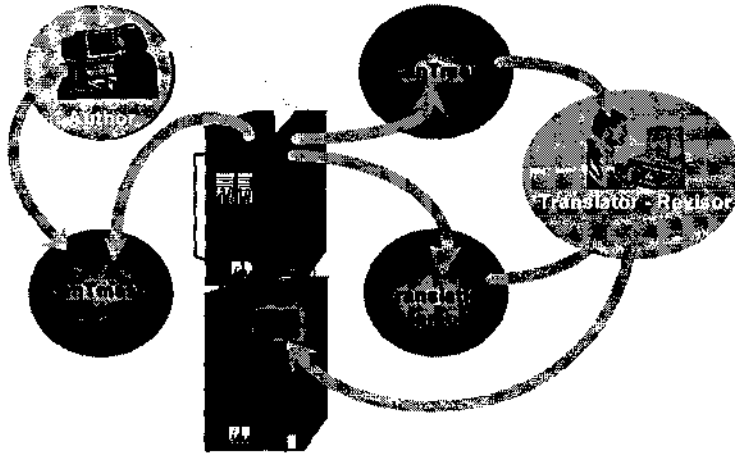
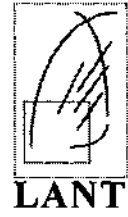
EAMT Copenhagen 97

# The Document Based Solution



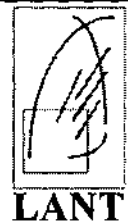
EAMT Copenhagen 97

# The Object Based Solution



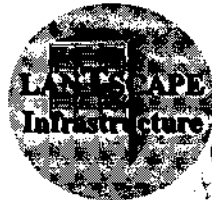
EAMT Copenhagen 97

# Multilingual Communication



## Translation Requests

E-mail  
Web-pages  
Attached Files



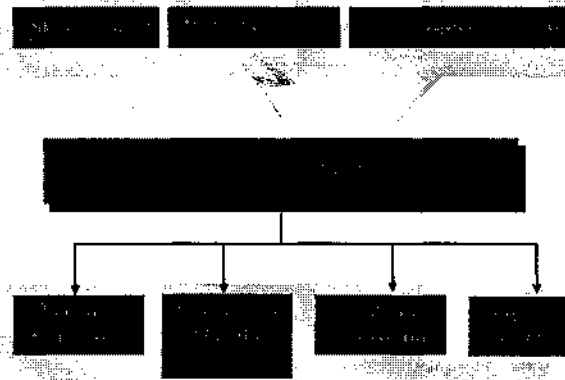
Machine Translation  
Systems



Translation Pool  
Agencies

EAMT Copenhagen 97

## LanTscape Architecture



EAMT Copenhagen 97

## Why to use Simplified English

- Increase of **Readability**
- Increase of **Understandability** for non-native speaker
- Increase of **Translatability**
- Decrease of **Liability risks**



EAMT Copenhagen 97

## Readability



LANT

### Non controlled :

Tighten the screw several turns, remove the screwdriver and check to see that the cover still is in place on the phone before connecting the phone wire.

### Controlled :

1. Tighten the screw several turns.
2. Remove the screw driver
3. Check to see that the cover still is in place on the phone
4. Connect the phone wire

EAMT Copenhagen 97

## Understandability



LANT

### Non controlled :

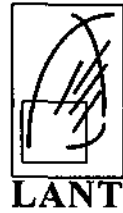
When fasteners are removed, always reinstall them at the same location from which they were removed.

### Controlled :

Always reinstall fasteners in the same location.

EAMT Copenhagen 97

## Translatability



### Non controlled :

A 10 megaohm input impedance digital voltmeter

### Controlled :

A digital voltmeter with 10 megaohm input impedance

EAMT Copenhagen 97

## Liability



### Non controlled :

Unscrew the plug-in unit from the wall and paint it.

### Controlled :

Unscrew the plug-in unit from the wall and paint the plug-in unit.

EAMT Copenhagen 97



## LANT<sup>®</sup>Master Grammar



- ✦ The system contains about 50 rules
- ✦ The rules coincide with the AECMA standard for Simplified English and they have been especially designed for translatability
- ✦ The rules keep a balance between generality and concreteness, specificity and redundancy
- ✦ Spelling errors and cases of ungrammatical input fall outside the rules in the grammar checker

EAMT Copenhagen 97

## System's Effect on the Usability of English and Translated Documents



- ✦ Full consistency in terminology use
- ✦ Tendency towards short sentences
- ✦ Tendency towards simplicity in syntactic structures
- ✦ High level of redundancy in syntactic structures
- ✦ Low level of verbosity on word level
- ✦ Low level of semantic complexity within sentences
- ✦ High consistency in spelling conventions
- ✦ Absence of ellipsis
- ✦ Absence of indirect rhetorical constructions
- ✦ Low level of ambiguity
- ✦ Tendency towards verbal, as opposed to nominal style
- ✦ Absence of pro-forms

EAMT Copenhagen 97

## LANT<sup>®</sup>Master Conformance Checker 1



- ✦ Interactive or batch mode
- ✦ The checker does not require strict conformance:
  - the user interface contains an overrule-error
  - the system deals with rule violation thresholds
  - the system only proposes autocorrections
- ✦ Nonconformances are communicated on three levels:
  - text-level
  - sentence-level
  - string-level

EAMT Copenhagen 97

## LANT<sup>®</sup>Master Conformance Checker 2



- ✦ The system provides help to the user in order to correct nonconformances:
  - all errors are diagnosed by means of a rule mark-up
  - all errors are located by marking the exact string in which the error occurs
  - all errors are autocorrected, in case the error can be automatically corrected by the system. The user can decide to accept the revision by the system, or not.

EAMT Copenhagen 97

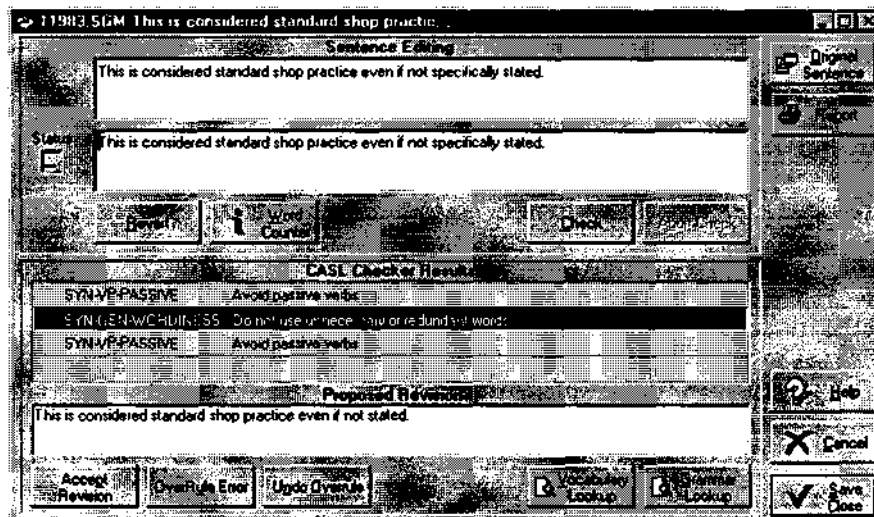
## LANT<sup>®</sup> Master Conformance Checker 3



- ✦ The following functionality should help the user to correct nonconformances:
  - a grammar help file which amply explains the rules and gives examples of good and bad use
  - a lexicon help file which contains glossaries of correct and incorrect word use
  - the interactive check mode which enables the user to immediately check the results of his editing
  - a file report functionality which enables the user to communicate with the system administrator

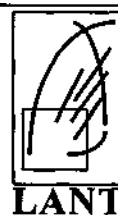
EAMT Copenhagen 97

## User Functionality



EAMT Copenhagen 97

## Technology Used



EAMT Copenhagen 97

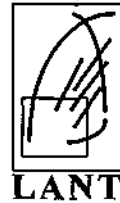
## Benefit 1

- ✦ Productivity increase
- ✦ Reduction of translation costs
- ✦ Reduction of time-to-market for dissemination of multilingual documentation
- ✦ Consistency of terminology throughout all documents
- ✦ Relieve of routine work for translators



EAMT Copenhagen 97

## **Benefit 2**



- ✦ High accuracy in translation
- ✦ Automatic preservation of format information
- ✦ Flexibility of work flow through batch processing

EAMT Copenhagen 97