INTRODUCTION
IPR is settled for all text sources
Text and tree alignment tools are released as open source software
Implementation of automatic alignment and Tree alignment is ongoing
All corpus material has been collected, converted, marked-up, tokenized, parsed, and aligned

1. TEXT SOURCES
- collected in
Google news (24k).
Atranos project (125k)
IMIX project (1k)
QA-system output
News headlines
Autocue-subtitle pairs
Essais "Book translations

2. TEXT PROCESSING
- tokenization
Conversion to visualizes
XML documents
- treebank
The hierarchical structure of
Syntactic nodes are aligned and labeled with a semantic similarity relations,
Erwin Marsi

3. TEXT SEGMENTATION
- contextual
- sentence
The text is automatically aligned down to the sentence level
Text segments in context.
The XML documents

4. TEXT ALIGNMENT
- automatic
Alignment is partly automatic
Next alignment is manually corrected using the new
Alignment is manually

5. SEMANTIC RELATION LABELING
- similarity
Five mutually exclusive semantic similarity relations

STATE OF AFFAIRS
On the basis of this corpus, developing software for automatic alignment and semantic relation labeling
Building

REFERENCES
Hendrickx, Jakub Zavrel
Deterding, Rijke
McCarthy, Hendrickx, Deterding, Rijke

For more information about the project, contact:

Erwin Marsi
Marsi

E: {Phone : + 31 - 13 - 4663070
The Netherlands
NL-5000 LE
Tilburg University
Erwin Krahmer
Contact information
Website: http://daeso.marsi.uvt.nl/

and Dementia
Parkison
and Dementia
three cups of coffee a day
daily coffee
risc
Dementia
restates
 specifies
chance
Dementia