

Software for processing patent claims

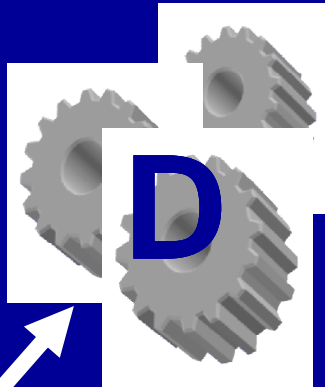
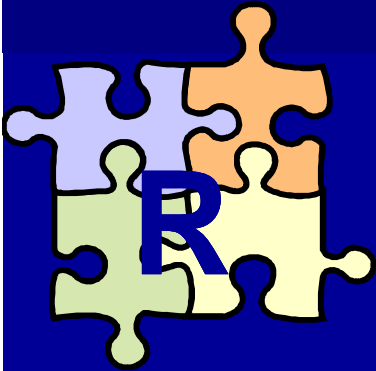
Svetlana Sheremetyeva
LanA Consulting ApS
Denmark

www.lanaconsult.com

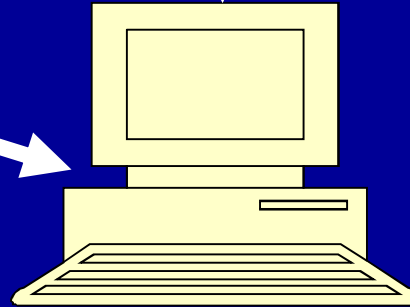
Overview

- *LanA Consulting ApS* - IP software
- Why develop special software for patents and patent claims
- AutoKey-P
 - key phrase extraction from patents and/or claims
- AutoTrans-C
 - Machine translation of patent claims

Lana Consulting



Patents
Claims



Patent & claim specific
Linguistic knowledge

Keywords

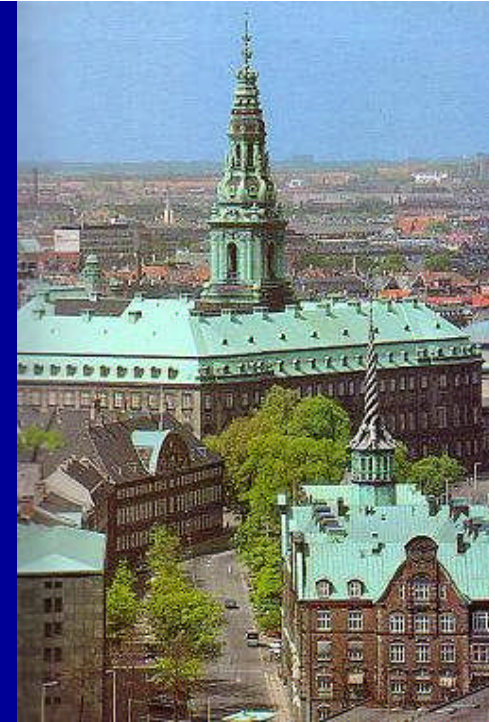
Claim trees

Claim features

Claim translation

Claim text

www.lanaconsult.com



Denmark
+45 51 22 34 58
lanaconsult@mail.dk

- TAEUS, MIT
- USA
- CrossLanguage
- Japan
- Zacco
- Scandinavia

Why develop special software for patents and patent claims

- Patent text differs greatly from general language
 - A lot of terminology & specific grammar
 - Long complex sentences
- Claim text differs greatly from patent description
 - One possibly extremely long sentence with extremely complex and *specific grammar* structure
 - Claims with different subject matters have different terminology and *grammar specificity*
- **No existing computer system**, can process patent texts (let alone claims) adequately due to their grammar complexity.

High quality processing
of patents and patent claims
is only possible when
patent and claims language
specificity is accounted for

AutoKey-P

Automatic extractor of high quality keywords and phrases

- Why do we need keywords?

Quick understanding, Search

- Where do we get them?

Clients, experts

- How ?

Manually, automatically

AutoKey-P

Automatic extractor of high quality keywords and phrases

- Do automatic keyword extractors exist?

Yes, rather good, normally statistical

- Why do it then?

To get better results

Better - Rather good = MONEY

- Why is **AutoKey-P** better?

Based on the combination of statistical & linguistic techniques & *specific linguistic knowledge* (lexicons and grammar) about patents and claims

www.lanaconsult.com

AutoKey-P

Automatic extractor of high quality keywords and phrases

In what way is **AutoKey-P** better?

- **Better quality:**

Grammatically correct noun phrases, closely related to the content, both for top and bottom relevancy lists

- **Quicker**

A patent is processed in a fraction of a second

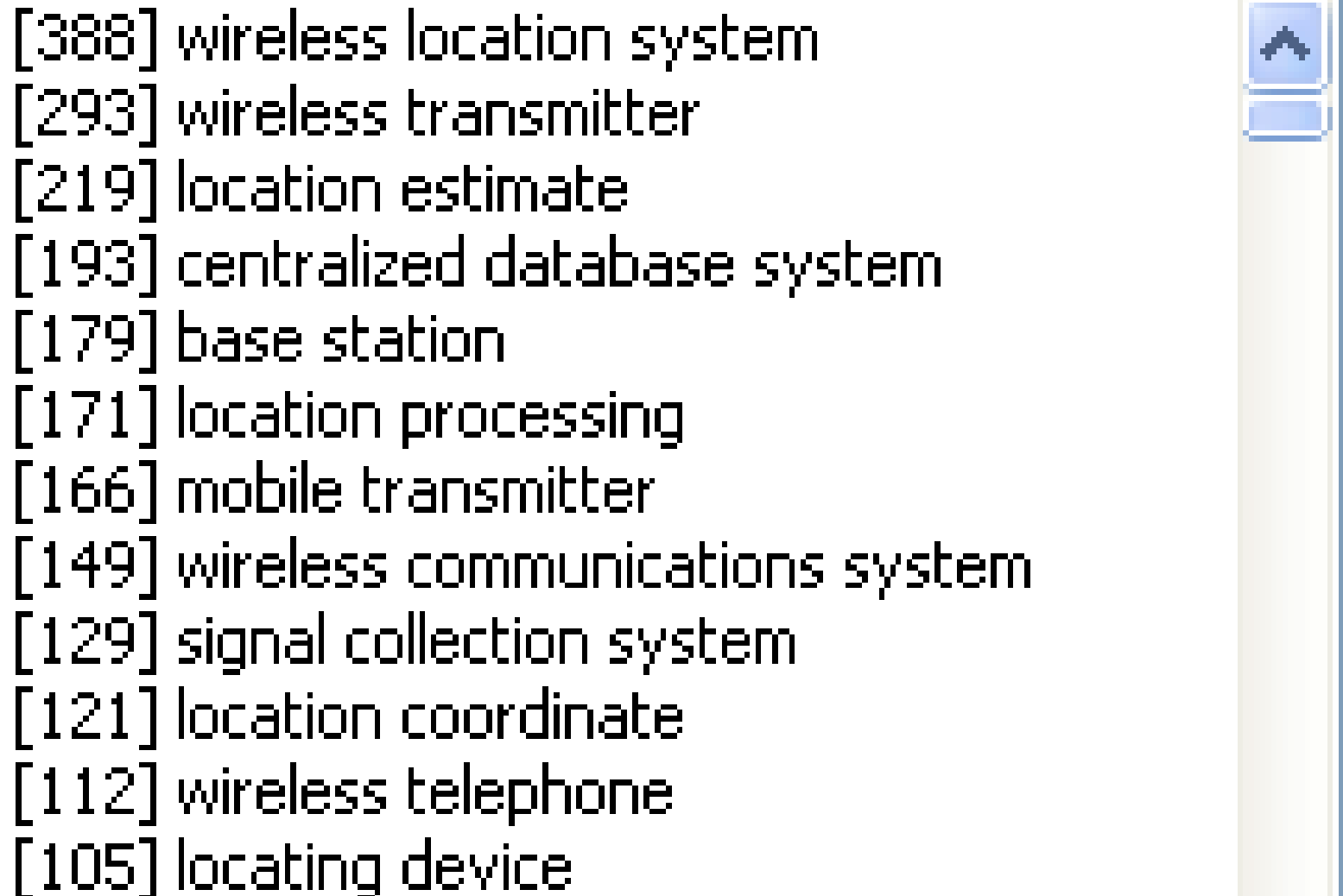
- **Wider range of application**

Human expertise

Automatic indexing of patent databases

Terminology acquisition for machine translation

A sample of **AutoKey-P** output (top relevancy)



[388] wireless location system
[293] wireless transmitter
[219] location estimate
[193] centralized database system
[179] base station
[171] location processing
[166] mobile transmitter
[149] wireless communications system
[129] signal collection system
[121] location coordinate
[112] wireless telephone
[105] locating device

A sample of **AutoKey-P** output (bottom relevancy)

- [2] overlay wideband signal
- [2] packet channel transmission
- [2] paging reception mode
- [2] paging signal
- [2] pcs frequency band
- [2] performance guardian process
- [2] periodic synchronization pulse
- [2] phase noise phase-locked loop
- [2] position signal
- [2] position-dependent characteristics of mob
- [2] potential location determination
- [2] power access channel transmission

AutoTrans-C

a machine translation system for patent claims

Prototype

- Based on claim specific linguistic knowledge (lexicon & grammar)
- High quality output + Automated postediting
- Development environment (tools)
 - It is possible to improve translation quality by updating MT rules and lexicon **without** extra programming
- Modular architecture
 - **High reusability**
 - Language rules are mostly the same across languages
 - **Foreign MT engines** can be merged into APTrans to speed up the development process

Every new pair of languages takes much less effort and development time

Why do we need a specific system for patent claim translation

- A patent claim is the most important part of a patent. It defines the scope of the invention and is the subject of legal protection.
 - Possibility of new procedure of international patenting (only claims are to be translated)
- **Translating patent claims is a more complex task, than translating patent disclosures**
 - They have a very complex syntactic structure, different from the rest of the patent disclosure
- **No existing computer system even developed for patents, can translate patent claims adequately.**

A sample claim text

What is claimed is: 1. A method for use in a wireless location system (WLS) that estimates the geographic location of a mobile station (MS), wherein the WLS overlays at least a portion of the geographic area of a wireless communications system, wherein the WLS uses radio and location processing resources to locate said MS using a process involving receiving a transmission from the MS at multiple sites and processing the received transmissions to estimate the location of the MS using time difference of arrival (TDOA) and/or angle of arrival (AOA) location processing techniques, comprising: passively monitoring communications between at least one base transceiver station (BTS) and at least one base station controller (BSC) of the wireless communications system; extracting MS information from the monitored communications, said extracted MS information including radio frequency channel information, whereby the WLS is able to receive transmissions from the MS; and using the extracted MS information to trigger location processing and to determine which radio frequency channel to use in performing location processing for said MS.

Why AutoTrans is feasible

- AutoTrans is tuned to a patent claim language (grammar) specificity
 - The language used in patent claims is complex but **very restricted** (no past or future tense, etc.)
 - If an MT system is to cover a restricted language subsystem the scope of development effort decreases greatly.
- Legal requirements for patent claims in Europe and North America are similar.
 - National claim sublanguages are much more close than national languages in general → It is possible to reuse a lot of language resources and programs and thus save on MT development when covering new language pairs.



From: English

To: Danish

Subject matter: Apparatus

Chunked Claim

- A self-centring support
- + for bearings
- comprising
- + two reversibly connected half-shells
- provided
- + with corresponding cavities
- adapted
- + to form a seat
- + for a bearing
- characterized in that
- at least one of the cavities
- is shaped
- so as
- + to form three radial raised portions
- + for the contact
- + of the bearing
- + along corresponding imaginary lines
- parallel
- + to the rotation axis
- + of said bearing

Sentences

Generic features of your invention:

A self-centring support for bearings COMPRISES two reversibly connected half-shells.

Two reversibly connected half-shells ARE PROVIDED with corresponding cavities.

Corresponding cavities ARE ADAPTED to form a seat for a bearing.

Novelty features of your invention:

At least one of the cavities ARE SHAPED to form three radial raised portions, for the contact and of the bearing.

Corresponding imaginary lines ARE PARALLEL to the rotation axis.

Workflow in APTrans

Source Language (SL)
Claim

Target Language Claim

SL ANALYZER

Template Transfer / Phrase MT

TL GENERATION

Result of Claim Decomposition

Short | Full | Original | Predicate Structure IN | Predicate Structure OUT | **Transfer** | Out text

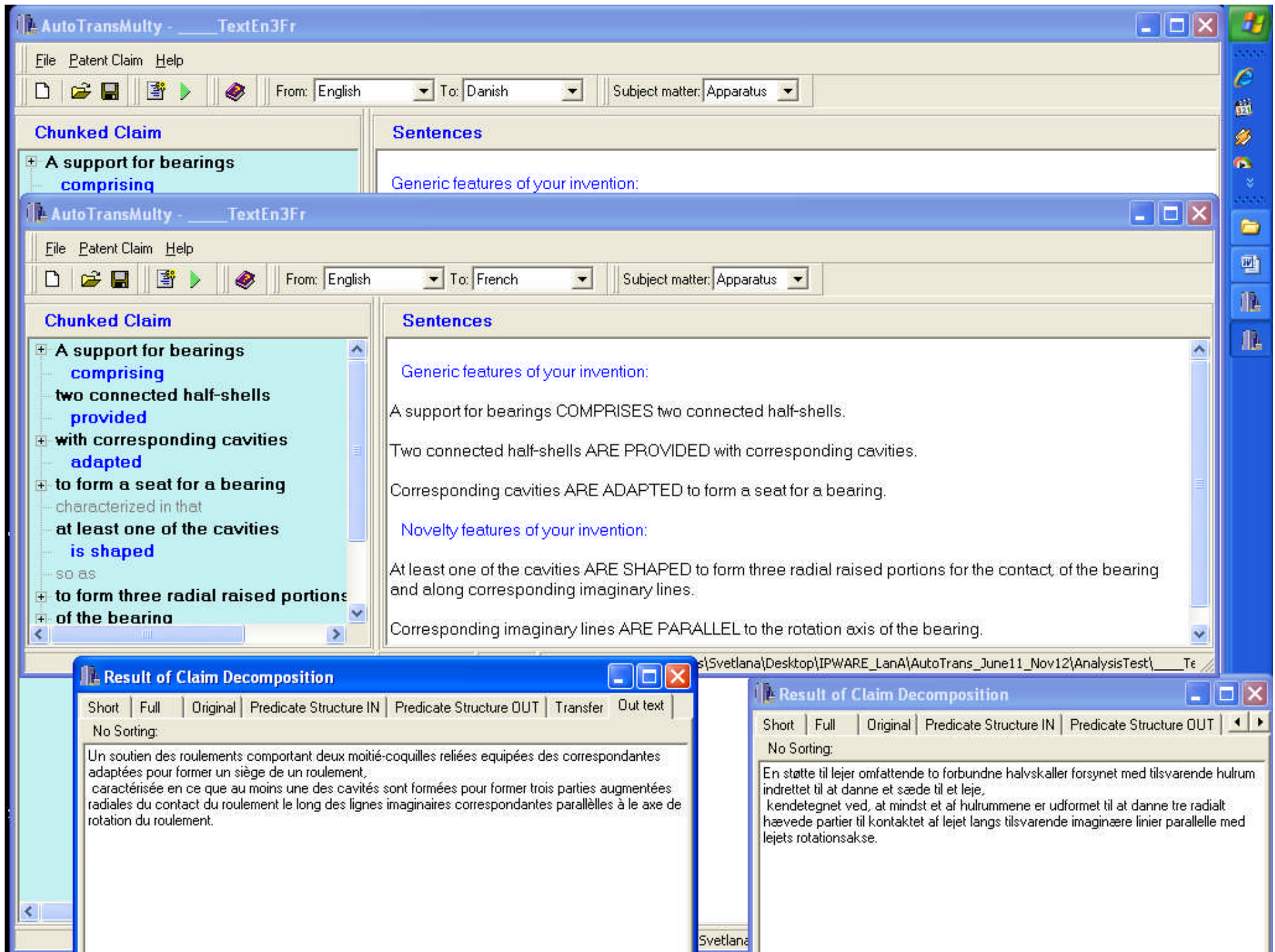
Before Transfer: After Transfer:

```
// SL Predicate structure
(P1 comprising
  1 A self-centring support for bearings
  2 two reversibly connected half-shells
)
(P2p provided
  1 two reversibly connected half-shells
  2 with corresponding cavities
)
(P3p adapted
  1 corresponding cavities
  2 to form a seat for a bearing
)
(P4p is shaped
  1 at least one of the cavities
  3 to form three radial raised portions for the contact
)
(P5p parallel
  1 corresponding imaginary lines
  2 to the rotation axis of said bearing
)
```

// TL Predicate structure after TRANSFER stage

```
(P1 omfattende
  1 en selvcentrerende støtte til lejer
  2 to reversibelt forbundne halvskaller
)
(P2p forsynet
  1 to reversibelt forbundne halvskaller
  2 med tilsvarende hulrum
)
(P3p indrettet
  1 tilsvarende hulrum
  2 et sæde til et leje
)
(P4p udformet
  1 mindst et af hulrummene
  3 til at danne tre adialt hævede partier til kontaktet af lejet
)
(P5p parallel
  1 tilsvarende imaginære linier
  2 med rotationsaksen af lejet
)
```

Save As... Close



Contact info

Svetlana Sheremetyeva
LanA Consulting ApS
lanaconsult@mail.dk

www.lanaconsult.com