

The Healthcare Sector and Language Technology

Lingsoft®

**LANGUAGE
MANAGEMENT**

*August 29, 2013
Michael Stormbom*



International full service
language company

Leading language
technology

Services designed for
analysis, processing and
utilisation of written and
spoken language





The Healthcare Sector

Challenges and
Requirements

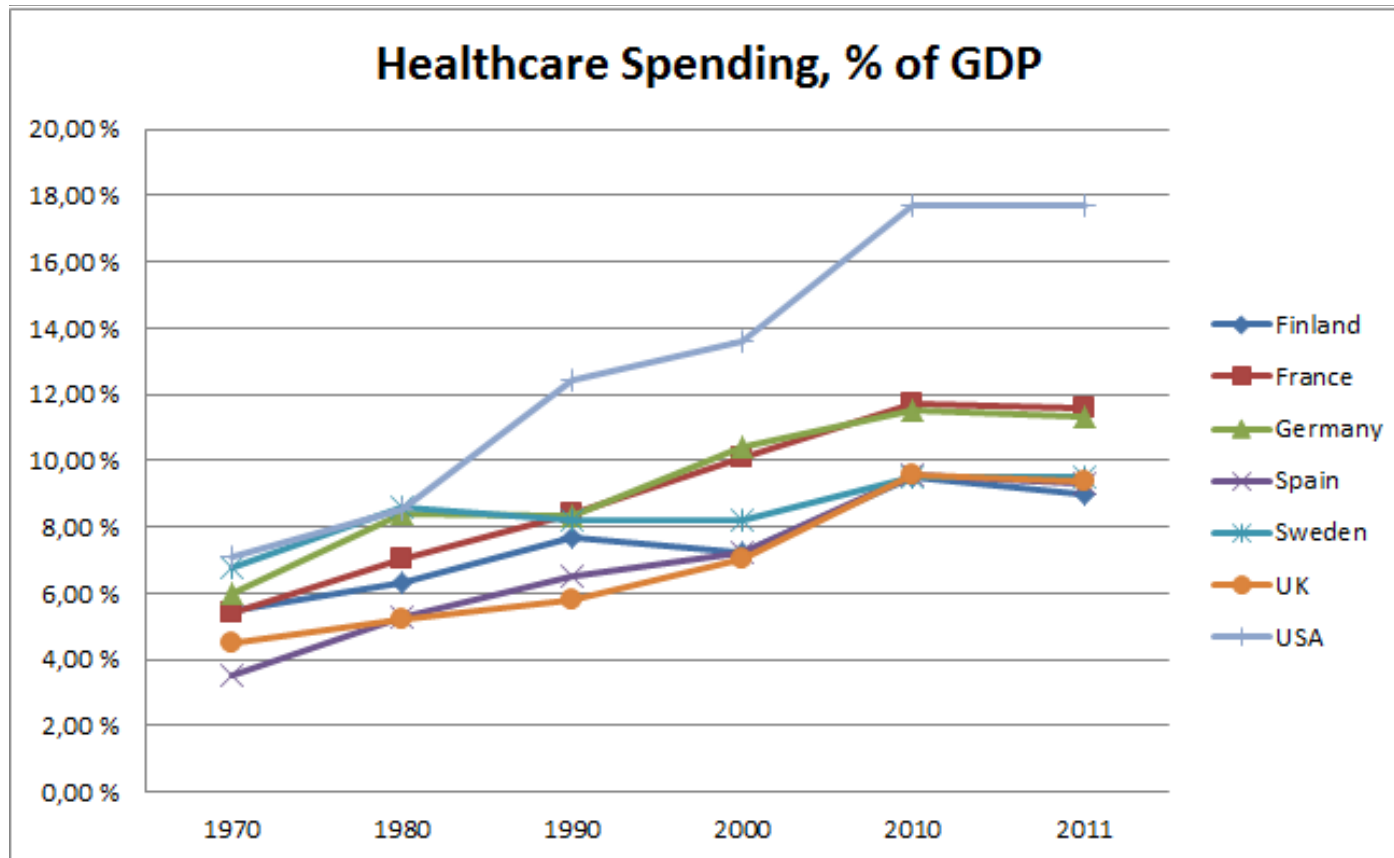
Medical Language

Case Study:
Medical Transcription and LT

The Healthcare Sector

- Healthcare sector one of the largest industries in the world
 - More than 10% of GDP in many developed nations
- More than 25 million people directly employed worldwide
- Various organisations, institutions and companies
 - Healthcare providers
 - Support and management
 - Administration
 - Technology
 - Pharmaceutical
 - ...

The Healthcare Sector



Spending Drivers

- Life expectancy
 - At the beginning of the 20th century: 30 years
 - 2010 average: 67 years
- Age distribution
- Lifestyle
 - Sedentary
 - Diet
 - Stress

- Human resources
 - Labour migration
 - Increasing needs for eg. nurses
- Patient safety
 - “To err is human, but to really screw up you need a computer”
 - Therac-25
- Cost control

- Multilingualism
 - Free mobility within EU
 - Many countries have several official languages
 - Patients, doctors, nurses may speak different languages
- Crossborder cooperation
 - Directive 2011/24/EU on patients' rights in cross-border healthcare
 - eHealth interoperability framework
 - Medical tourism
- Translation and interpretation needs

- Cross-organisation data sharing needs
- Standardisation
 - A large set of different standards
 - Various stages of adoption
 - Legacy data
 - **Terminologies and classification**

- **ICD-10**

- **Diagnosis classification maintained by WHO**
- **Hierarchical structure**
- ***“B30.0+H19.2 Keratoconjunctivitis due to adenovirus”***
- **ICD-11 exp. 2015**

- **ATC**

- **Drug classification maintained by WHO**
- ***“A01AA02 sodium monophluorophosphate”***

Layman Comprehension

- Patients have the rights to see their records
- Difficult to comprehend without significant domain knowledge
- **Translating very domain-specific and incomprehensible language into “plain English”**
- **Simultaneous translation into multiple languages**

- **In Finland (pop. 5.4 mil) on a yearly basis:**
 - **Six million polyclinic visits**
 - **One million hospital visits**
 - **One million private healthcare visits**
 - **All of which are to be documented**
- **Mostly in Finnish, occasionally in Swedish**
- **Patient journals**
- **Compactness and speed**
- **Standards (ICD-10)**

- **Finnish**
 - **Finno-ugric language**
 - **Agglutinative**
 - **Extensive case system**
 - **Complex morphology**
 - **Rich in inflections**
- **Swedish**
 - **Germanic language**
 - **Limited inflections**

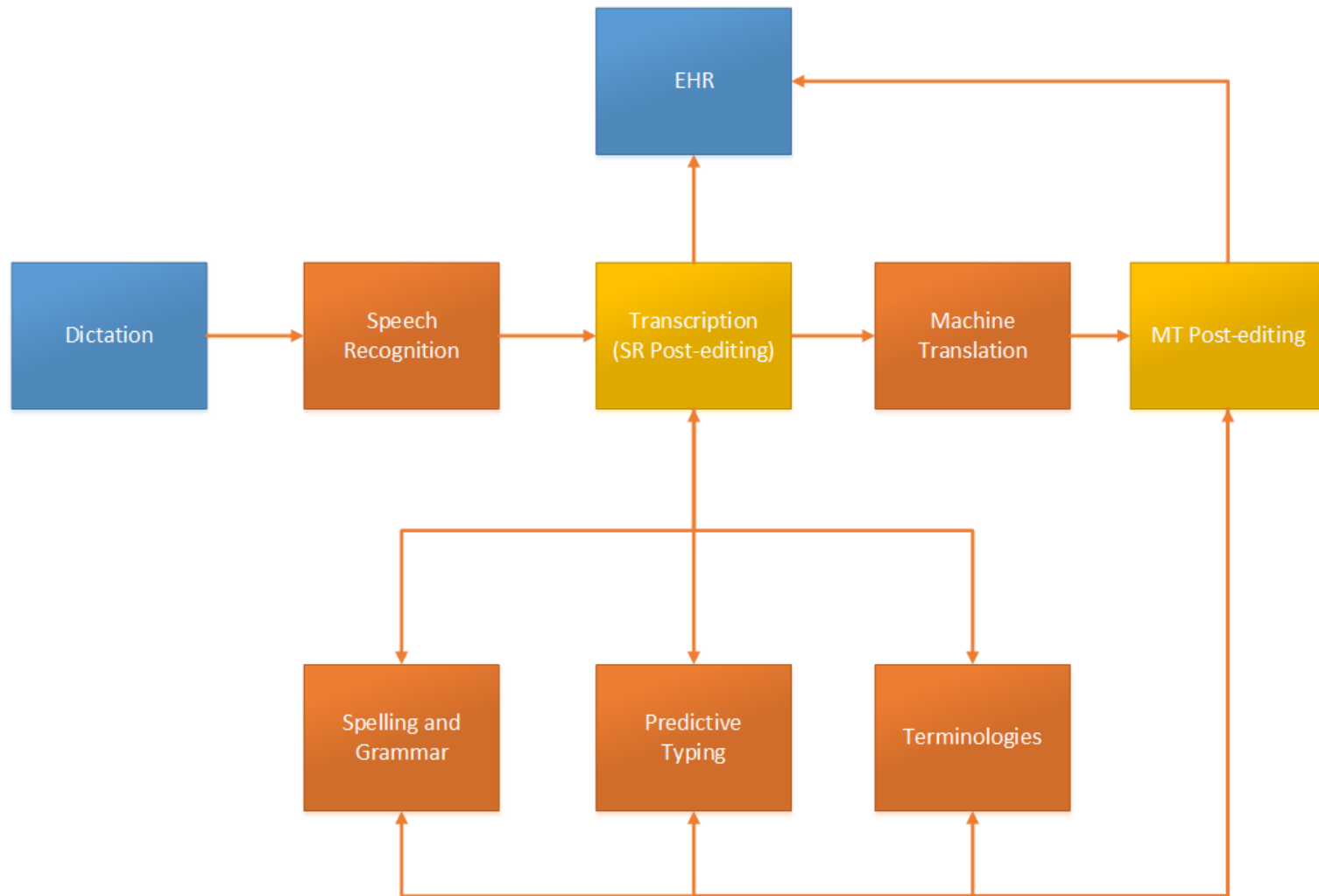
- Finite verbs and subjects absent
- Finnish: *Verenpaine alhainen. [The patient's] bloodpressure [is] low.*
Swedish: *Lågt blodtryck. [The patient] [has] low bloodpressure.*
- Non-informative parts omitted
Swedish: *det* with impersonal verbs (*det regnar, it rains*)
- *Joku on tehnyt jotakin → Joku tehnyt jotakin*
Someone has done something → Someone done something

- **Acronyms and numerical data in abundance**
- **Conjunctions replaced with commas**
- **Independent main clauses joined together with commas**
- **Latin and native terminology used interchangeably**

- **Finnish:** *Antikoagulaatiohoito aloitettu.*
Swedish: *Antikoagulationsbehandlingen inledd.*
The anti-coagulation treatment [has been] started.
- **Finnish:** Free word order
 - *Pöydällä on kirja. There is a book on the table.*
 - *Kirja on pöydällä. The book is on the table.*
 - Interchangeable in medical text
- **Subordinate clauses infrequent**

- **By leveraging language technology, workflows and process can be optimised**
- **Different LT used in different parts of the workflow**
- **Reducing manual labour**
- **Allowing LSPs to provide more efficient and less labour-intensive services**
 - **Larger volumes and lower unit cost**

Medical Transcription Workflow



Medical Transcription Workflow

- **Feedback loop**
 - **Coverage continuously improved**
 - **Terminologies to speller and predictive typing**
 - **Supporting morphological and syntactical analysis**
- **The role of the transcriber**
 - **Transcription**
 - **Post-editing and correction**
 - **Feedback**

- **Medical domain language model**
- **Recording quality and devices**
 - **Hospital infrastructure**
- **Pronunciation**
 - **Non-native speakers**
- **Post-processing (capitalization, punctuation, numbers in the right case)**

Terminology Management

- **Terminology checks**
- **Lookup and research tool**
 - **Eg. ICD-10 codes**
 - **Synonyms**
- **Multiword expressions**
- **Multilingual terminologies**
- **Central repository for vocabularies**

Spell and Grammar Checking

- Correct language includes correct spelling and grammar
 - Deliberately ungrammatical
- Context-sensitive spell checking
Their is a house
- False positives
 - Domain vocabulary
- Predictive typing for eg. ICD-10

- Level of urgency
 - Translations may be needed rapidly at any time during the day
- Accuracy
 - Exact information needs to be conveyed
- With human post-editing or verification
- **GF Demo**

- Standard RGLs
 - Ungrammatical structures
- Pre- and postprocessing
 - Punctuation, abbreviation consistency
 - Spelling errors in source
- Lexical resources, medical FIN and SWE
 - ICD-10, ATC
- Iterative approach

- The healthcare sector is facing many challenges
- Leveraging LT will be an important aspect in overcoming those challenges
- In particular multilingual LT is becoming increasingly important



<http://www.oecd.org/health/health-systems/oecdhealthdata.htm>

<http://www.corepointhealth.com/sites/default/files/whitepapers/hl7-history-v2-v3.pdf>

<http://www.who.int/classifications/icd/en/>

<http://sydney.edu.au/medicine/fmrc/atc/>

http://www.iso.org/iso/products/standards/catalogue_ics_browse.htm?ICS1=35&ICS2=240&ICS3=80&

Vainikainen, Kirsi, Klinisk minilekt i elektroniska patientjournaler, University of Turku 2010

Thank you!

www.lingsoft.fi
info@lingsoft.fi
[@lingsoft](#)
slideshare.net/lingsoft

Lingsoft®
**LANGUAGE
MANAGEMENT**

