Medical Frames as Target and Tool

Lars Borin, Maria Toporowska Gronostaj, Dimitrios Kokkinakis
Göteborg University
Department of Swedish Language
Språkdata/Språkbanken
Sweden
{first.last}@svenska.gu.se

Abstract

In this paper we present a pilot study on the development of a FrameNet-like annotation of a sample of Swedish medical corpora, for a selected set of verbal predicates. We explore and exploit a number of linguistic tools for the provision of much of the necessary annotations required by such a semantic scheme. Particular attention is paid to the syntactic and semantic roles of scheme elements. We discuss in detail methodological issues and take up the relevance of our research for natural language processing (NLP) tasks.

1 Introduction

The conviction that enrichment of corpora with annotation layers of syntactic and semantic information will provide valuable support for refined text mining has been the main impetus for this corpus oriented pilot study. We have explored cumulative morphosyntactic text processing as a preliminary stage in semantic tagging. The main goal of our study has been to examine whether such integration of information can in a significant way contribute to semi-automatic acquisition and extraction of semantic schemes from corpora, in particular in the medical domain. By semantic schemes we mean frame-like constructions analogous to those in FrameNet. Formally, "FrameNet annotations are constellations of triples that make up the frame element realization for each annotated sentence" (Ruppenhofer et al., 2006:6), i.e. grammatical function [e.g. Subject]; frame-element [e.g. HUMAN]; phrase type [e.g. NP]. FrameNet resources have been recently developed for a number of languages, e.g., Spanish, German and Japanese. The FrameNet project (Baker et al., 1998) builds upon the theory of semantic frames formulated by Fillmore (1976), supported by corpus evidence. It is assumed here that access to such formalized semantic schemes can significantly improve the semantic component of a number of NLP tasks requiring semantic processing, including question-answering, automatic semantic role labelling, natural language generation, and information extraction (IE), in which there is a direct correspondence between frame-like structures and templates. Templates in the context of IE are frame-like structures with slots representing the basic components of events (cf. Surdeanu et al., 2003).

Related work is presented in section 2. The methodology underlying the morphological, syntactic and semantic pre-processing is outlined in section 3. Section 4 deals with the issues concerning lexical annotation of medical corpora. In section 5 we discuss the possibility of semi-automatic acquisition of frames based on qualitative and quantitative criteria. We end the article with conclusions and discussion.

2 Related Work

There are a number of approaches to FrameNet-like annotation including the influential work by Gildea and Jurafsky (2002) and Gildea and Palmer (2002), who point to the necessity of using syntactic information for the semantic annotation task and for predicting semantic roles based on the FrameNet corpus; the use of named-entity recognition by Pradhan et al. (2004) and others; see for instance the CONLL 2004 and CONLL 2005 shared tasks for semantic role labeling[1] and the SemEval-2007 Frame semantic structure extraction

[1] <http://www.lsi.upc.edu/ ~srlconll/>

task.[2] In our context, the work by Johansson & Nugues (2006) on Swedish is of particular relevance. In their work a corpus was annotated using cross-language transfer from English to Swedish. However, closer to our goals has been the work described by Wattarujeekrit et al. (2004); Huang et al. (2005), Cohen and Hunter (2006) and Chou et al. (2006) within the (bio)medical domain.

3 Methodology

3.1 Corpus Sampling and Annotation

We started by sampling a large number of sentences from the MEDLEX corpus (Kokkinakis, 2006), a large collection of articles from the medical domain, currently comprising about 45,000 documents. The sampling was performed after the identification and selection of a set of 30 important verbs, according to their significance compared to general newspaper corpora and which indicate events containing medical entities. Examples of such verbs are *operera* ‘to operate’, *behandla* ‘to treat’, *injicera* ‘to inject’, *vaccinera* ‘to vaccinate’ and *palpera* ‘to palpate. The medical entities were supplied by the use of a Swedish MeSH tagger[3] for the categories *anatomy* (A), *organisms* (B), *diseases* (C), *chemicals and drugs* (D), *analytical, diagnostic and therapeutic techniques and equipment* (E), and *psychiatry and psychology* (F). Although MeSH is a valuable resource, it is rather limited in coverage considering the wealth of terminology in medical language. Therefore, we have complemented the MeSH annotations by developing a module that recognizes important types of (medical) terms, particularly *names of pharmaceutical products, drugs, symptoms* and (anatomical) *Greek and Latin terms*. Named entity tags were also added to the sample. A generic named entity tagger was applied which recognizes and annotates eight main types of named entities; *person, location, organization, object/artifact, event, work, time* and *measure expressions*; for details see Kokkinakis 2004.

[2] <http://framenet.icsi.berkeley.edu/semeval/FSSE.html>

[3] The Medical Subject Headings (MeSH) is the controlled vocabulary thesaurus of the U.S. National Library of Medicine (NLM), widely used for indexing medical data. The MeSH is a hierarchical thesaurus. The Swedish MeSH tagger is based on the Swedish translation made by staff at the Karolinska Institute Library (<http://mesh.kib.ki.se/swemesh/>) which contains 22,325 entries. MeSH is the central vocabulary component of the UMLS, frequently used as a provider of lexical medical information for biomedical natural language processing tasks (bio-NLP).

The net effect of the preprocessing described in this section is that the NPs in the sample sentences are annotated with their semantic classes, which turns out to be a very useful piece of information to have when parsing the sentences.

3.2 Streamlining Parsing with Semantic Classes

Grammatical functions are one of the main features and prerequisites for the realization of FrameNet annotations. Therefore, the semantic class annotations described above, together with part-of-speech tags, were merged into a single representation format and fed into the syntactic analysis module, which is based on the Cass parser (*Cascaded analysis of syntactic structure*; see Abney 1997). The Cass parser is capable of annotating grammatical functions and is designed for use with large amounts of (noisy) text. Cass uses a finite-state cascade mechanism and internal transducers for inserting actions and roles into patterns. The Swedish grammar used by the parser has been developed by Kokkinakis and Johansson Kokkinakis (1999), and has been modified and adapted in such a way that it is aware of the features provided by the pre-processors, particularly the medical terminology.

The annotations produced by the entity and terminology taggers significantly reduce the complexity of the sentence content, which in turn reduces the complexity of the parsing task, since the sentences contain fewer tokens, with less complex phrases, and thus can be more reliably parsed. Consider the example in figure 1, which, after the pre-processing stages, has been reduced from 26 to 10 tokens and 6 annotations, while a complex noun phrase, *cancer coli Duke’s B*, has been replaced by a single label, ‘<DISEASE>’.

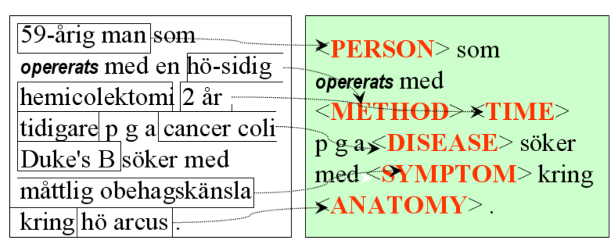


Figure 1. Simplification of input sentences

The syntactic analyses produced by the parser were in turn transformed into the TIGER-XML interchange format (König & Lezius, 2003), a flexible graph-based architecture for storage, indexing and querying of syntactically analyzed texts (appendix

1a). Our main purpose for doing this was that we wanted to apply existing software for manual frame annotation and for the analysis and inspection of the results, namely the SALSA/ SALTO tool (Burchardt *et al.*, 2006), which requires TIGER-XML input, thus minimizing the software development overhead (appendix 1b). Using this method we are now in the process of developing a semantically annotated sample that can be further used for experiments with machine learning algorithms.

4 Medical Frames as Target

4.1 Medical Frames in FrameNet

Access to multilayered lexical and grammatical information representing the content of texts is one of the prerequisites for an efficient understanding and generation of natural language. The FrameNet approach, with roots in Fillmore's case roles, offers an interesting approach to the study of lexical meaning described in terms of semantic frames. Semantic frames are generalisations of conceptual scenarios evoked by predicates and their frame elements. According to Ruppenhofer et al. (2006) there are roughly 780 semantically related frames (10,000 word senses/lexical units) accounted for in FrameNet. For each frame, there is a set of lexical units listed and exemplified with semantically and syntactically tagged examples from the British National Corpus (BNC). A small subset of these frames pertain directly to medical scenarios, like *Medical conditions, Experience bodily harm, Cure, Health response, Recovery, Institutionalization, Medical instrument.* Other, more general ones like *Placing* and *Removing,* do this in an indirect way by including lexical units of medical terminology dealing with notions of implanting or removing body parts. An overview of a repository of medically related frames in FrameNet with specification of core and non-core frame elements is provided in appendix 2b. The core frame elements, capturing the semantic valence of predicates, are obligatory ones, while the non-core ones add optional information.

The semantic salience of the types of core elements listed in appendix 2b applies also to Swedish. However, whenever designing frame-like schemes for specific sub-domains, further descriptive detail might be called for. Conflation of conceptually similar frame elements, e.g. Ailment and Affliction, semantic role overlap between general and specific roles as for example Agent and Healer, and postulation of new medical schemes are some of the issues which need to be considered when building a similar resource with focus on medical scenarios for Swedish.

4.2 From Frame Elements to MeSH Categories and Scheme Elements

Mapping medical frame elements onto the corresponding concepts in a thesaurus-based lexicon turns a relatively information-poor lexical resource into a more expressive and robust one and hence more useful for semi-automatic semantic annotation of corpora. For annotating the Swedish corpus, we have used our thematically sorted lexicons with medical vocabulary and the Swedish data from MeSH.

Since the MeSH vocabulary is sub-classified according to topics like anatomy, diseases etc., there is a possibility of mapping between some medical core concepts in the FrameNet and the top nodes in MeSH classification including their hyponyms. The results of this mapping are indicated in table 1:

Core frame elements in FrameNet	**MESH thesauristic nodes**
Ailment, Affliction	Diseases
Body_parts	Anatomy
Medication	Chemicals and Drugs
Treatment	Analytical, Diagnostic and Therapeutic Techniques and Equipment
Patient	Persons

Table 1. Mapping core frame elements onto MeSH top nodes

As already mentioned above (section 3.1), the tag set based on the MeSH top nodes has been further enlarged with thematic lists for both medical concepts like *symptoms* and supplementary named entities such as *time*, *location*, *measure* etc. All of these occur frequently in combination with the verbs selected for our sample. Since the sample came from a medical corpus, the instantiated uses of the verbs represent predominantly their medical senses. To make the semantic medical schemes appear more distinct the corpus sentences have been syntactically pre-processed, i.e., complex syntactic phrases containing syntactic dependences have been analysed to find their semantic heads, which have been subjected to semantic annotation, with the exception of noun phrases containing two or more medical tags. The latter will undergo further analysis for detecting types of medical collocations. Examples (i) and (ii) below illustrate the annotated corpus.

(i) *<TIME> har <PERSON-GRP> opererats i <PLACE> för sina <SYMPTOM> i <ANATOMY> . (Original sentence: "Sedan 1987 har cirka 7 000 personer opererats i Sverige för sina svettningsproblem i händerna")*

(ii) *<DISEASE> i <ANATOMY> - <DISEASE> - kan <TIME> opereras med utmärkt resultat om durationen är <TIME> . (Original sentence: "Bristning i centrala retina - makulahål - kan idag opereras med utmärkt resultat om durationen är under 4-6 månader .")*

As follows from the above, the focus in our work is on the semantic types of referents, and thus our methodology contrasts with the FrameNet approach which takes the predicate and the evoked role scenario as the point of departure for determining a set of frame elements. The tags in our corpus are meant to provide a first approximation of medical semantic schemes by naming the types of annotated elements. To make the distinction between FrameNet and our approach clear, the terms *semantic schemes* and *scheme elements* are used henceforth in our study. A quantitative overview of semantic tags in the sample sentences (700 000 tokens) is given in the table 2.

Semantic labels	# in the whole sample (# with *operera*)
DISEASE	22 100 (1 346)
ANATOMY	11 080 (1 528)
CHEMICAL	10 450 (186)
METHOD	2 276 (467)
ORGANISM	4 090 (7)
PERSON	12 434 (1460)
PERSON-GRP	11 810 (829)
LOCATION	3 024 (216)
TIME	19 131 (897)
MEASURE	3 732 (319)

Table 2. Semantic annotations in the sample sentences

4.3 Case Study: Medical Senses of *operera* 'to operate'

To assess the correctness of our assumptions and the possible advantages or disadvantages of the chosen methodology, we have taken a closer look at the Swedish verb *operera*, whose medical sense ('perform surgery') is not described in FrameNet. The verb *operera* is polysemous in both Swedish and English, but only its medical senses are considered below, as the corpus and the pilot study is restricted to the medical sub-domain. In the following we select some of the frequent schemes instantiated in the corpus in order to examine the types of the medical scenarios this verb can evoke (appendix 1c illustrates dependency concordances with *operera*). The verb *operera* in its medical readings occurs in the corpus as either a simplex, reflexive or particle verb (phrasal verb) followed by the particles *bort* or *ut* (away, out) or *in* (in), as illustrated below:

- **simplex** *operera*: two sub-senses and thus two partly different schemes are represented in the corpus:

(i) to give consent to and undergo a surgical procedure with PERSON used in the double role of both semi-Agent and Experiencer, with ANATOMY and DISEASE as possible core arguments;

*e.g. <PERSON (semi-Agent & Experiencer)> **har** precis **opererat** <ANATOMY> i <ANATOMY> (Original sentence: Jag **har** precis **opererat** min laterala menisk i vänster knä)*

(ii) to perform a surgical procedure, with one PERSON in the role of Patient, another PERSON in the role of Agent (Medical professional), DISEASE and BODY PART as possible core arguments

*e.g. <PERSON(Patient)> **opererades** <TIME> av <PERSON (Agent)> (Original sentence: Han **opererades** omedelbart av dr Piotr)*

*<PERSON (Patient)> som är <MEASURE> **har** både strålats och **opererats** för <DISEASE> (Original sentence: "min pappa som är 63 **har** både strålats och **opererats** för tonsillscancer")*

- **reflexive** *operera sig*: to give consent to have a surgical procedure performed with PERSON in the double role of semi-Agent and Experiencer and DISEASE

*e.g. <PERSON (Experiencer)> **har opererat mig** för <DISEASE> i <ANATOMY> som var <MEASURE> (Original sentence: Jag **har opererat mig** för malignt melanom i ryggen som var 1,2 mm)*

- **particle verb** with two sub-senses:

(i) to give consent to removing or implanting a body part or an implant with semi-Agent & Experiencer and ANATOMY or IMPLANT as possible scheme elements.

*e.g. <PERSON (semiAgent & Experiencer)> **opererade bort** <ANATOMY> för <TIME> (Original sentence: "Jag **opererade bort** blindtarmen för ganska exakt 36 timmar sedan")*

(ii) to perform a surgical procedure aiming at removing or implanting a body part or an implant with PERSON in role of Agent (medical professional), ANATOMY, IMPLANT and optionally with PERSON being a Donor as

possible scheme elements. IMPLANT and Donor have not been annotated in the examined corpus. (The tag IMPLANT will be reserved for an artefacts, since organic implants are tagged as ANATOMY.)
e.g. <PERSON (Agent)> ***opererade in*** *en p-stav i den kvinnliga <ANATOMY> (Original sentence: "Läkaren* ***opererade in*** *en p-stav i den kvinnliga patientens arm")*

This specification of scheme elements captures some prototypical scenarios for the verb *operera*. The schemes can undergo certain modifications resulting in null instantiation of scheme elements, which can be either constructional, definite or indefinite (Fillmore et al. 2003).

5 Semi-automatic Acquisition of Semantic Schemes

Semi-automatic acquisition of semantic schemes on the basis of an annotated corpus is far from a trivial task for verbs such as *operera*, mainly due to the fact that the human subject, when used in active form can correspond to different semantic roles, ranging from the agentive ones, e.g. Agent usually manifested by medical professionals to a semi-agentive in Experiencer role and non-agentive in the Patient role. The question remains whether there are explicit supportive cues to distinguish between those role instances and whether other roles can be semi-automatically tagged. Some proposals which might be worth testing with respect to role identification for the examined verbs are:

Agent: Medical professional

- lexical criterion: checking the list of lexical units naming medical professionals;
- presence of a prepositional phrase introduced by *av* followed by a scheme element PERSON in a sentence in passive voice;
- presence of another np in the same scheme labelled as PERSON (Patient).

Experiencer:

- presence of a noun annotated as PERSON in a scheme and an inalienable noun annotated with the label ANATOMY having either a definite form (*Jag opererade bort blindtarmen)* or preceded by a possessive pronoun referring to the subject (*Jag har precis opererat min laterala menisk [...])*;
- reflexive use of the verb (*Jag har opererat mig för malignt melanom).*

Patient:

- presence of an explicit Agent in the same scheme;
- presence of an implicit Agent in the same scheme (passive voice);
- object in an active sentence or subject in the passive sentence annotated with the tag PERSON.

Anatomy:

- lexical criterion: checking an available sub-lexicon.

Disease:

- lexical criterion: checking an available sub-lexicon;
- syntactic cue: use of preposition *för* in construction *operera någon för DISEASE* (*cf.* English *operate on sb (for sth)*)

For a preliminary listing of schemes for the analysed verb senses see appendix 2a.

6 Conclusions

The advantages of the pre-processing and the consequences for lexical annotation have been illustrated and we believe that given the results of our case studies, the described methodology represents a feasible way to proceed in order to aid the annotation of large textual samples. As advantages of lexical annotation, the following needs mentioning:

- relevant semantic schemes can be retrieved from medical corpora
- integrated layers of syntactic and semantic annotation support the acquisition of semantic roles and thus enhance text understanding
- the semantic schemes provide input for various NLP tasks
- semantically annotated nouns promote disambiguation of predicates
- access to semantic schemes can support classification of lexical units carrying related meaning (e.g. *operera bort, avlägsna, ta bort*)

The quantitative analysis of the examined corpus has shown that the importance of many linguistically optional scheme elements needs to be reassessed when viewed from a medical pragmatic perspective. For example Time, Measure and Method provide relevant data for diagnosing patients' health condition. Another issue that may need special attention in future annotating tasks is that of tagging pronouns. It seems that these should not be tagged before anaphoric relations and their semantic roles have been established. This is particularly important for distinguishing between patients and health care providers. The figures in table 2 illustrate clearly the importance of identifying and annotating different entity types, particularly for the annotation of FrameNet non-core elements such as Time, Measure and Method, but also a strong indication of the frequency of important core elements such as Disease and Anatomy.

References

S. Abney. 1997. Part-of-speech tagging and partial parsing. *Corpus-Based Methods in Language and Speech Processing*. S. Young and G. Bloothooft (eds), 118–136. Kluwer AP.

C.F. Baker, C.J. Fillmore, and J.B. Lowe. 1998. The Berkeley FrameNet Project. *Proc. of the 36th Annual Meeting of the ACL and the 17th International Conference on Computational Linguistics (COLING-ACL 1998)*. Montreal.

A. Burchardt, K. Erk, A. Frank, A. Kowalski, and S. Pado. 2006. SALTO - A versatile multi-level annotation tool. *Proceedings of the 5th International Conference on Language Resources and Evaluation*. Genoa, Italy.

W.-C. Chou, R.T. Tsai, Y.-S. Su, W. Ku, T.Y. Sung, and W.-L. Hsu. 2006. A semi-automatic method for annotating a biomedical proposition bank. *Proc. of the Workshop on Frontiers in Linguistically Annotated Corpora.* 5–12. Sydney, Australia

K.B. Cohen and L. Hunter. 2006. A critical review of Pasbio's argument structures for biomedical verbs. *BMC Bioinformatics*

C.J. Fillmore. 1976. Frame semantics and the nature of language. *Annals of the NY Acad. of Sciences Conference on the Origin and Development of Lnguage and Speech*. Vol. 280.

C.J. Fillmore, C.S. Johnson, and M.R.L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16(3).

D. Gildea and D. Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

D. Gildea and M. Palmer. 2002. The necessity of parsing for predicate argument recognition. *Proc. of ACL 2002*, Philadelphia, PA.

M. Huang, X. Zhu, and M. Li. 2005. A hybrid method for relation extraction from biomedical literature. *Journal of Medical Informatics*.

R. Johansson and P. Nugues. 2006. A FrameNet-based semantic role labeler for Swedish. *Proc. of Coling/ACL 2006*. Sydney, Australia

D. Kokkinakis. 2004. Reducing the effect of name explosion. *Proc. of the Beyond Named Entity Recognition, Semantic Labelling for NLP Tasks*. workshop at LREC. Lisbon, Portugal

D. Kokkinakis. 2006. Collection, encoding and linguistic processing of a Swedish medical corpus – The MEDLEX experience. *Proc. of the 5th LREC*. Italy.

D. Kokkinakis and S. Johansson Kokkinakis. 1999. A cascaded finite-state parser for syntactic analysis of Swedish. *Proc. of the 9th European Chapter of the Association of Computational Linguistics (EACL)*. Bergen, Norway.

E. König and W. Lezius. 2003. The TIGER language – A description language for syntax graphs, Formal definition. Technical report. Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.

S. Pradhan, W. Ward., K. Hacioglu., J. Martin, and D. Jurafsky. 2004. Shallow semantic parsing using support vector machines. *Proc. of the Human Language Technology Conference/North American chapter of the ACL (HLT/NAACL)*, Boston, MA.

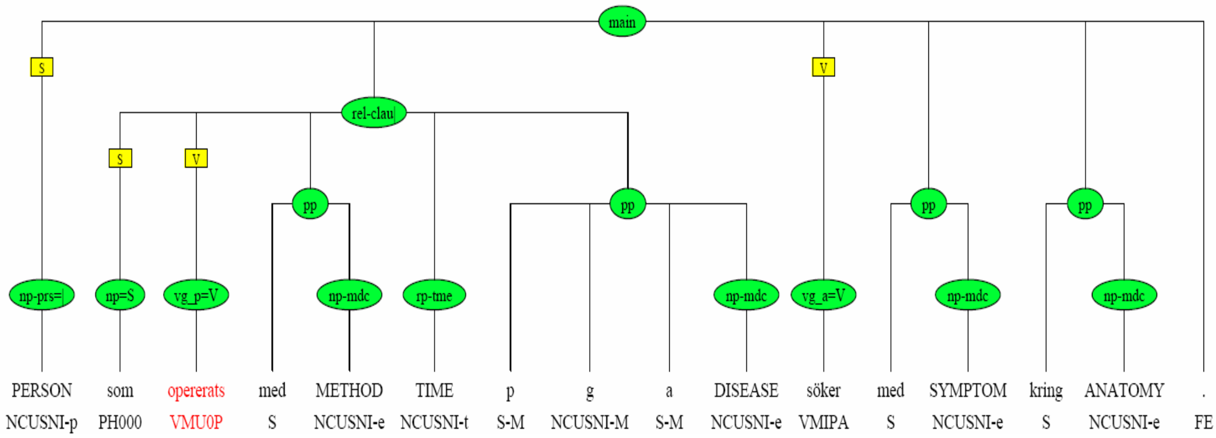
J. Ruppenhofer, M. Ellsworth, M.R.L. Petruck, C.R. Johnson, and J. Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice.* Available from <http://framenet.icsi.berkeley.edu/book/book.pdf>

M. Surdeanu, S. Harabagiu, J. Williams, and P. Aarseth. 2003. Using predicate-argument structures for information extraction. *Proc. of the 41st Annual Meeting of the Assoc. of Comp. Ling,* 8–15.

T. Wattarujeekrit, P.K. Shah, and N. Collier. 2004. Pasbio: Predicate-argument structures for event extraction in molecular biology. *BMC Bioinformatics* 2004, 5:155

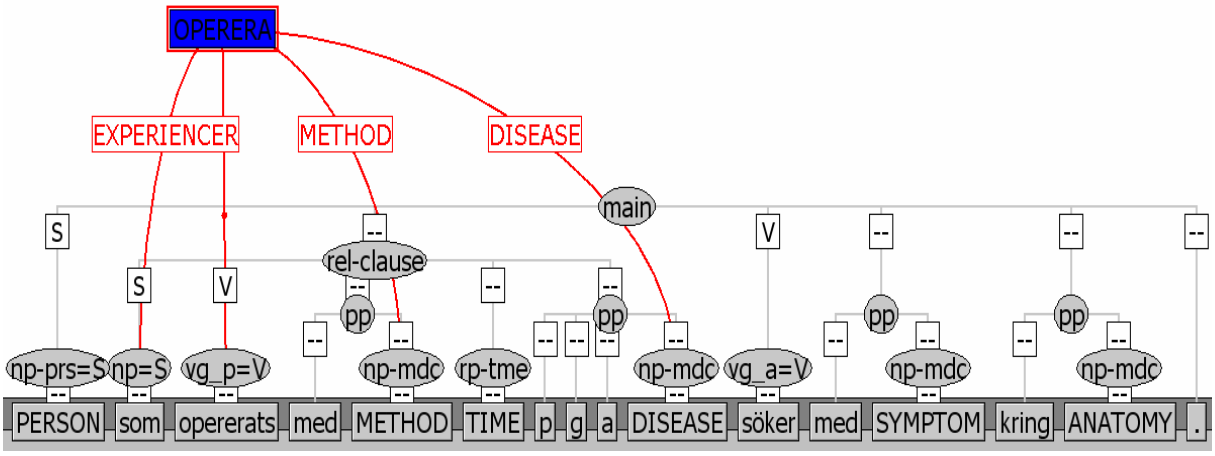
Appendix

1a



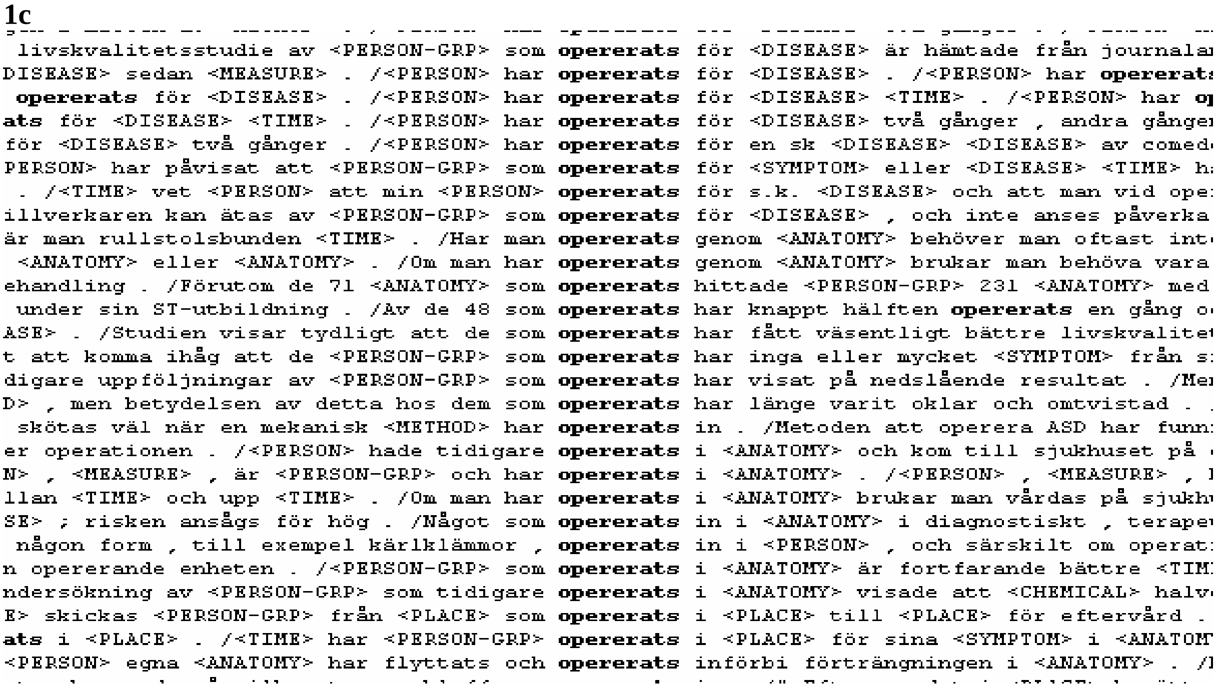
Syntactic analysis

1b



Role Assignment

1c


livskvalitetsstudie av <PERSON-GRP> som **opererats** för <DISEASE> är hämtade från journala
DISEASE> sedan <MEASURE> . /<PERSON> har **opererats** för <DISEASE> . /<PERSON> har **opererats**
opererats för <DISEASE> . /<PERSON> har **opererats** för <DISEASE> <TIME> . /<PERSON> har **o**
ats för <DISEASE> <TIME> . /<PERSON> har **opererats** för <DISEASE> två gånger , andra gånge
för <DISEASE> två gånger . /<PERSON> har **opererats** för en sk <DISEASE> <DISEASE> av comed
PERSON> har påvisat att <PERSON-GRP> som **opererats** för <SYMPTOM> eller <DISEASE> <TIME> h
. /<TIME> vet <PERSON> att min <PERSON> **opererats** för s.k. <DISEASE> och att man vid ope
illverkaren kan ätas av <PERSON-GRP> som **opererats** för <DISEASE> , och inte anses påverka
är man rullstolsbunden <TIME> . /Har man **opererats** genom <ANATOMY> behöver man oftast int
<ANATOMY> eller <ANATOMY> . /Om man har **opererats** genom <ANATOMY> brukar man behöva vara
ehandling . /Förutom de 71 <ANATOMY> som **opererats** hittade <PERSON-GRP> 231 <ANATOMY> med
under sin ST-utbildning . /Av de 48 som **opererats** har knappt hälften **opererats** en gång o
ASE> . /Studien visar tydligt att de som **opererats** har fått väsentligt bättre livskvalite
t att komma ihåg att de <PERSON-GRP> som **opererats** har inga eller mycket <SYMPTOM> från s
digare uppföljningar av <PERSON-GRP> som **opererats** har visat på nedslående resultat . /Me
D> , men betydelsen av detta hos dem som **opererats** har länge varit oklar och omtvistad .
skötas väl när en mekanisk <METHOD> har **opererats** in . /Metoden att operera ASD har fun
er operationen . /<PERSON> hade tidigare **opererats** i <ANATOMY> och kom till sjukhuset på
N> , <MEASURE> , är <PERSON-GRP> och har **opererats** i <ANATOMY> . /<PERSON> , <MEASURE> ,
llan <TIME> och upp <TIME> . /Om man har **opererats** i <ANATOMY> brukar man vårdas på sjukh
SE> ; risken ansågs för hög . /Något som **opererats** in i <ANATOMY> i diagnostiskt , terape
någon form , till exempel kärlklämmor , **opererats** in i <PERSON> , och särskilt om operat
n opererande enheten . /<PERSON-GRP> som **opererats** i <ANATOMY> är fortfarande bättre <TIM
ndersökning av <PERSON-GRP> som tidigare **opererats** i <ANATOMY> visade att <CHEMICAL> halv
E> skickas <PERSON-GRP> från <PLACE> som **opererats** i <PLACE> till <PLACE> för eftervård .
ats i <PLACE> . /<TIME> har <PERSON-GRP> **opererats** i <PLACE> för sina <SYMPTOM> i <ANATOM
<PERSON> egna <ANATOMY> har flyttats och **opererats** införbi förträngningen i <ANATOMY> . /

Semantic Concordance

Appendix

2a

Scheme: V *operera*	**Exempel**
PERSON(Agent) **V** PERSON(Patient)	Vi har opererat två patienter med Budd-Chiaris syndrom; Även kirurgen som opererat henne tog sig tid för att deltaga
PERSON(Agent) **V** (an instance of indefinite null instantiation)	I dagsläget opererar fyra urologer vid hans klinik; När läkarna opererar, suger slangarna blodceller genom lasern
PERSON(Agent) **V** METHOD	Roboten opererar med fyra armar
PERSON(Agent) **V** DISEASE	De opererar aldrig näsfrakturer
PERSON(Agent) **V in/ut** IMPLANT	Oftast opererar man in en mekanisk klaffprotes Risken för ett nytt benbrott finns alltid när man opererar ut metallimplantatet
PERSON(Agent) **V bort/ut** ANATOMY	Man opererar bort hela njuren,
PERSON(Agent)**V bort** ORGANISM	När man opererar en pinoidalcysta
PERSON(semi-Agent&Experiencer) **V** ANATOMY	Jag har precis opererat min laterala menisk i vänster knä
PERSON(semi-Agent&Experiencer) **V sig för** DISEASE	Jag har opererat mig för malignt melanom i ryggen

Schemas for the verb *operera*

2b

Frame	**Core frame elements**	**Non-core frame elements**
Medical_conditions	Ailment, Patient	Body_part, Cause, Degree, Name, Symptom
Experience_bodily_harm	Body part, Experiencer	Containing_event, Duration, Frequency, Injuring_entity, Iterations, Manner, Place, Severity, Time
Cure	Affliction, Body_part, Healer, Medication, Patient, Treatment	Degree, Duration, Manner, Motivation, Place, Purpose, Time
Health_response	Protagonist, Trigger	Body_part, Degree, Manner
Institutionaliztion	Authority, Facility, Patient	Affliction, Depictive, Duration_of_final state, Explanation, Manner, Means, Place, Purpose, Time
Recovery	Affliction, Body part, Patient,	Company, Degree, Manner, Means,
Medical_instruments	Instrument	Purpose
Medical_professionals	Professional	Affliction, Age, Body _system, Compensation, Contract_basis, Employer, Ethnicity, Origin, Place_of_employment, Rank, Type
Medical specialties	Specialty	Affliction, Body_system, Type
Observable_bodyparts	Body_part, Possessor	Attachment; Descriptor, Orientational_location, Subregion,
Placing	Agent, Cause, Theme, Goal	Area, Beneficiary, Cotheme, Degree, Depictive, Distance, Duration, Manner, Means, Path, Place, Purpose, Reason, Result, Source, Speed, Time
Removing	Agent, Cause, Source, Theme	Cotheme, Degree, Distance, Goal, Manner, Means, Path, Place, Result, Time, Vehicle

Medical frames in FrameNet

NODALIDA-2007

FRAME 2007: Building Frame Semantics Resources for Scandinavian and Baltic Languages

Workshop organizers:
Pierre Nugues and Richard Johansson
Department of Computer Science, Lund University

LUND UNIVERSITY

Department of Computer Science
http://nlp.cs.lth.se

Institute of Computer Science
http://math.ut.ee