

# Characteristics and Analysis of Finnish and Swedish Clinical Intensive Care Nursing Narratives

Helen Allvin<sup>f</sup>, Elin Carlsson<sup>f</sup>, Hercules Dalianis<sup>f</sup>, Riitta Danielsson-Ojala<sup>a</sup>,  
Vidas Daudaravičius<sup>b</sup>, Martin Hassel<sup>f</sup>, Dimitrios Kokkinakis<sup>c</sup>, Heljä Lund-  
gren-Laine<sup>a</sup>, Gunnar Nilsson<sup>f</sup>, Øystein Nytrø<sup>d</sup>, Sanna Salanterä<sup>a</sup>, Maria  
Skeppstedt<sup>f</sup>, Hanna Suominen<sup>e</sup>, Sumithra Velupillai<sup>f</sup>

<sup>a</sup>Department of Nursing Science, University of Turku, VSSHP, Turku, Finland,

<sup>b</sup>Vytautas Magnus University, Lithuania,

<sup>c</sup>Department of Swedish, University of Gothenburg, Sweden,

<sup>d</sup>IDI, The Norwegian University of Science and Technology, Norway,

<sup>e</sup>NICTA Canberra Research Laboratory and Australian National University, Australia

<sup>f</sup>Department of Computer and Systems Sciences/Stockholm University

Forum 100

SE-164 40 Kista, Sweden

<http://www.dsv.su.se/hexanord>

## Abstract

We present a comparative study of Finnish and Swedish free-text nursing narratives from intensive care. Although the two languages are linguistically very dissimilar, our hypothesis is that there are similarities that are important and interesting from a language technology point of view. This may have implications when building tools to support producing and using health care documentation. We perform a comparative qualitative analysis based on structure and content, as well as a comparative quantitative analysis on Finnish and Swedish Intensive Care Unit (ICU) nursing narratives. Our findings are that ICU nursing narratives in Finland and Sweden have many properties in common, but that many of these are challenging when it comes to developing language technology tools.

## 1 Introduction

The purpose of this study<sup>1</sup> is to do content and lexical analysis of nursing narratives written in an

Intensive Care Unit (ICU). The ultimate goal of our research is to define linguistic similarities and language-specific aspects that differentiate clinical narratives in Finnish and Swedish in order to lay groundwork for developing internationally applicable language technology solutions and create a framework for characterising and comparing clinical narratives. Free text is handy for information entry but a challenge for information extraction, care handover and other uses of gathered information. Language technology can alleviate some of these problems in retrospective analysis by offering a more semantically informed interpretation and abstraction. However, the most promising potential of language technology is to interactively improve, interpret and code during text entry so that the resulting structured, coded, free text can be validated immediately. The critical bottleneck today is namely information handover and reuse, and extensive text is simply not used nor is useful. Interactively validated, semantically processed text could be more usable and support abstraction, visualization and query tools for the benefit of clinicians, patients, researchers and quality administrators.

---

<sup>1</sup> Our research on the Stockholm EPR Corpus (Dalianis et al., 2009) has been approved by Etikprövningsnämnden i Stockholm, (the Regional Vetting Board), reference number

---

2009/1742-31/5. Our research on the Finnish Corpus (Salanterä et al., 2009) has been approved by the Ethical committee of the Hospital District of South West Finland, reference number 2/2009, §66.

In this paper, we analyze Finnish and Swedish ICU nursing narratives from both qualitative and quantitative perspectives. Our data includes textual nursing documentation of adult patients with a protracted inpatient period. We have chosen ICUs because of their international similarity in decision making (Lauri & Salanterä 2002) and nursing documentation because it covers the entire inpatient period.

## 2 Background

### 2.1 Clinical text

*Clinical text* covers the text documents produced for clinical work by clinicians and occurs in clinical information systems. It is written by *clinicians*, that is, professionals (physicians, nurses, therapists and other specialist) responsible for patient care. Its primary purpose is to serve patient care as a summary or hand-over note. However, clinical text is also written for legal requirements, care continuity and purposes of reimbursement, management and research. Clinical text covers every care phase and, depending on the purpose, documents differ. Documents that describe the patient's state, current health problems and socio-medical history are very different from those describing a care plan, its actualization and evaluation of care outcomes. Again, these differ from diagnostic notes, lab results, radiography readings, pathology reports and discharge documents that plan further care at discharge. Finally, clinical text may have been entered in "real time", in retrospect, or as a summary, by the bedside or elsewhere. The enterer can be a clinician, secretary who transcribes a dictate, speech recognition software or another system that generates or synthesizes text, (McDonald 1997, Thoroddsen et al., 2009.).

### 2.2 Legal requirements for clinical documentation in different countries

In several countries clinical documentation is based on law. In Finland, the Ministry of Social Affairs and Health (Statutes of Finland, 298/2009) defines that to ensure good care, all necessary and wide-ranging information has to be registered in patient records. In Sweden, the National Board of Health and Welfare has a similar approach (Patientdatalagen 2008:355). Clinical text should be

explicit and intelligible, and only generally well-known, accepted concepts and abbreviations are allowed to be used. It should detail adequately the patient's conditions, care and recovery.

### 2.3 Special features of ICU and nursing

An ICU is an essential component of most large hospitals with high quality care. ICUs provide care for critically ill patients and focus on conditions that are life-threatening and require comprehensive care and constant monitoring (Webster's 2010). This task is fairly similar universally. It is based on optional, international guidelines focusing on triage, admission, discharge and education. This international similarity was evident when nurses' decision making was studied in Canada, Finland, Northern Ireland, Switzerland, Norway and the USA (Lauri & Salanterä 2002); the study showed that decision making of ICU nurses was the most uniform in different countries when compared with nurses working in public health care, psychiatric care, and short and long term care.

Clinical text written by nurses, that is, *nursing narratives*, both in Finland and in Sweden is based on the care process which stands for gathering information from the patient, setting goals for care, implementing nursing interventions, and evaluating the results of given care. In Finland, the national standardized documentation model has been implemented with the Finnish care classification (assessment, interventions and outcomes of care) (Tanttu & Ikonen 2007). The Swedish VIPS model provides a structure for the documentation process with key words that reflect the nursing process (Ehrenberg et al., 1996).

ICU nursing narratives can be lengthy, especially when the patient stay in the ICU is prolonged. As much as 60 A4 pages equivalents of written text may be gathered during one period of care. However, clinicians have somewhat different opinion on how to organize the information they write. For example, headings are often inconsistent and text under headings can cover a lot of other issues than those directly concerning the given heading. (Suominen et al., 2009.)

### 2.4 Related studies

Since most of the available clinical documents are in free-text form, a number of stylistically oriented efforts to characterize the data from various angles

have taken place. This may include various topics, from viewing detailed information about specific items (e.g. readability, Kim et al., 2007) to identifying patterns and structures in order to provide better technology to automatically process the sublanguage (Pakhomov et al., 2006). The majority of such efforts investigate different aspects of linguistic features at a monolingual level, for instance, Hahn & Wermter (2004); Tomanek et al., (2007); Chung (2009); Harkema et al., (2009); while for a thorough review of various related issues see Meystre et al., (2008). In the Nordic context, Josefsson (1999) discusses Swedish clinical language and shows examples on how verb constructions in a clinical setting differ from a non clinical setting. One claim is that the physician unmarks the verb forms for agentivity when writing about the patient and what actions she takes, for example, *Patienten hallucinerar* [*The patient hallucinates*] instead of the normal form *Patienten får hallucinationer* [*The patient experiences hallucinations*].

Hellesø (2005) describes nurses' general use of the language function in the nursing discharge notes. She finds that the text in the nursing discharge notes is information-dense and characterized by technical terms, and that the use of standardised templates helped nurses improve the completeness, structure and content of the information. Comparisons at a monolingual level between written clinical text and lay text has been carried out by Dalianis et al., (2009). A contrastive computational linguistics study was carried out between the Stockholm EPR Corpus (SEPR) and a general language corpus, both written Swedish text. The findings showed that SEPR contained longer words and that the vocabulary was highly domain-specific. Other work is described in Ownby, (2005). Comparing clinical text at a crosslingual level has, to our knowledge, only been done by Borin et al., (2007).

### 3 Analysis of Finnish and Swedish ICU nursing narratives

The analyzed nursing narratives origin from one ICU in a university-affiliated hospital both in Finland and Sweden. Our inclusion criterion was an ICU inpatient period of at least 5 days and patient's age of at least 16 years. The Finnish data includes nursing narratives from 514 patient records (496

unique patients, 18 rebounds, a patient record is defined as each inpatient period of at least 5 days per patient) between January 2005 and August 2006. The Swedish data includes nursing narratives from 379 patient records (333 unique patients, 46 rebounds) between January 2006 and May 2008. Since we did not have complete admission and discharge documents from both countries, our analysis is performed on daily nursing narratives. These documents are written by ICU nurses during the actual inpatient period from the patient admission to the discharge.

#### 3.1 Qualitative analysis

A manual content analysis was performed by four health care professionals (i.e., three native Finnish speakers knowing Swedish, one Swedish native speaker) and one native Swedish speaking language consultant. Three average-sized patient records each from Finland and Sweden were chosen for our analysis (average size 2,389 words for Finland and 5,169 words for Sweden). In the analysis, we considered special features (Table 1) of daily notes both from the structural and content related points of view.

The style and context of both Finnish and Swedish text is very similar. For health care professionals, and especially with an ICU background, all the texts are intelligible and the meaning of a writer becomes evident from the context even in the presence of numerous linguistic and grammatical mistakes; almost all the sentences are lacking both grammatical subjects and objects. It is evident that in both countries, the narratives are written from a professional to a professional in order to support information transfer, remind about important facts, and supplement numerical data.

A feature common for all the six records is that they rarely contain any subjects or objects when nurses are writing about patients. However, in the Swedish nursing narratives the word *patient* is used as a subject or object much more often than in the Finnish narratives. The abbreviation *pat.* is mostly used for this reference and *she/he* is never used for this purpose. In the whole data, *pat.* is 40 percent more common than *she/he*, which is the most common personal pronoun. It seems that the word *patient* or *pat.* is used more when the professionals are writing about relatives. In general, pronouns are used infrequently in the narratives, and

Special features of Finnish narratives		Special features of Swedish narratives	
Structure	Examples	Structure	Examples
Headings are used in 2 out of 3 patient records. Headings are typically used as subjects or subjects are partially used.	<i>Diuresis: occasionally profuse.</i> ( <i>Diureesi: ajoittain runsasta.</i> ) <i>Pupils move under eyelids but does not open eyes.</i> ( <i>Pupillit liikkuvat luomien alla, mutta ei avaa silmiään.</i> )	Headings are used in all daily narratives. In Swedish daily narratives, the structure of headings seems to be obligatory. The headings are used typically as subjects.	<i>Circulation: Stable with inotrop.</i> ( <i>Cirkulation: Stabil med inotropi.</i> ) <i>Reacts only for pain stimulation during the suction of intubation tube.</i> ( <i>Reagerar enbart vid smärtstimuli vid sugning i tuben.</i> )
Present and past participles are typical but verbs of <i>be</i> , <i>is</i> and <i>are</i> are not used.	<i>Consciousness remained unchanged.</i> ( <i>Tajunta pysynyt ennallaan.</i> ) <i>Blood pressure low.</i> ( <i>Verenpaine matala.</i> )	Present and past participles are typical but verbs of <i>be</i> , <i>is</i> and <i>are</i> are not used.	<i>Breathing: Ventilator parameters unchanged.</i> ( <i>Andning: Ventilator parametrarna oförändrade.</i> )
Complete sentences are rarely used.	<i>No spontaneous movements, rigidifies.</i> ( <i>Ei spontaania liikettä, jäykistele.</i> ) <i>Husband and daughter have been staying a long time beside the patient.</i> ( <i>Mies ja tytär olleet pitkään potilaan vierellä.</i> )	Complete sentences are rarely used.	<i>Light sedation, looks up now and then.</i> ( <i>Lätt sederad, tittar upp ibland.</i> ) <i>She took the wedding ring and the watch home.</i> ( <i>Hon tog med sig vigselring och klocka hem.</i> )
Misspellings are found but the content or meaning is still clear.	<i>Hemodynamic – hemodynamic</i> ( <i>Hemodynamiikka – henodynamiikka</i> )	Misspellings are found but the content or meaning is still clear.	<i>The mother is informed.</i> ( <i>Mammans är informered.</i> ) <i>Magnesium is added.</i> ( <i>Magnesium har tillsatts.</i> )
Content	Examples	Content	Examples
The word <i>patient</i> as a subject or object is infrequently mentioned. If this word is mentioned it is not abbreviated.	<i>Oxidates well or ventilates well.</i> ( <i>Happeutuu hyvin tai ventiloituu hyvin.</i> )	The word <i>patient</i> is used more often than in Finnish narratives as a subject or object. It is also replaced with abbreviations of <i>Pat</i> or <i>Pt</i> .	<i>Patient got a percutaneous tracheostomy today.</i> ( <i>Patienten har fått en perkutan trakeostomi idag.</i> ) <i>Very worried about patient's condition.</i> ( <i>Mycket oroliga över patientens tillstånd.</i> ) <i>Pt. wakes up for talking and appears to be adequate.</i> ( <i>Pt. vakner på tilltal och upplevs som adekvat.</i> )
Signs are typically used: e.g., >, <, -->, +, -.	<i>The height for the drain raised from 10 --&gt; 20 mmHg.</i> ( <i>Dreneerausrajaa nostettu 10 --&gt; 20 mmHg.</i> ) <i>Got medicine --&gt; good response.</i> ( <i>Sai lääkettä --&gt;hyvä vaste.</i> )	Many different abbreviations are used. The origin of entire word is Swedish, English, Latin, professional or ICU typical.	<i>em. [eftermiddag, afternoon], HR [heart rate], VF [Ventricula/Fibrillation, Ventricular Fibrillation]</i>

Table 1. Special structural and contextual features of Finnish and Swedish daily ICU nursing narratives. The original examples are added in ().

*I* very rarely. If the reader is not a health care professional, a risk for confusing the subject (i.e., the patient or nurse) arises. However, the context makes it almost always clear who is referred to. Approximately half of the narratives do not contain any verb. The most common tense is perfect, but

without the auxiliary *has*. When the meaning does not contain a subject it becomes “unnatural” to use *has*. Instead, the supine form is used, for example *slept*, *lain*, and *eaten*. Both present and past participles without *be*-verb are common, for example, *Breathing: Ventilator parameters unchanged*.

The use of headings is frequent and good – most of the time the content matches the headings (Tables 1 and 3). In addition, headings are used similarly in the Swedish and Finnish documents. Most of the time the headings are considered as subjects of the sentence, for example, *Consciousness: Unchanged. Liquor brighter than yesterday.*

However, in the use of headings there are two interesting findings: If the headings are to be chosen freely, as in the Finnish narratives, nurses tend to use their own headings and hence many synonyms or closely related concepts are used; for example, *hemodynamics* versus *blood pressure and pulse* or *breathing* versus *oxidation*. If the headings are obligatory, as in the Swedish narratives, nurses tend to write their observations under the heading which is somehow closest to the subject; for example, *body temperature* under *circulation* or *level of sedation* under *sleep*.

For both languages the use of different abbreviations is very common. Almost every daily nursing narrative included several abbreviations. Most of the abbreviations are typical for an ICU domain: *CVP* [*central venous pressure*], *PEEP* [*Positive End-Expiratory Pressure*], *EN* [*Enteral Nutrition*], *TPN* [*Total Parenteral Nutrition*], *pO2* [*partial pressure of oxygen*], *pCO2* [*partial pressure of carbon dioxide*], *MV* [*Minute Ventilation*] and *MAP* [*Mean Arterial Pressure*]. From a language technology point of view this means that ICU nursing narratives contain language-independent vocabulary. However, nurses in both countries also use many language dependent abbreviations.

### 3.2 Quantitative analysis

The Finnish data set (n=514) was quantitatively analyzed using the morphological analyser FinT-WOL and the disambiguator FinCG, (Lingsoft 2010), and the Swedish data set (n=379) using the GTA, Granska Text Analyzer (Knutsson et al., 2003). Both data sets are rich in terms of amount of text and vocabulary (Table 2). It is also clear that the amount of text written per day and patient varies a lot in both data sets. More complex words were spelled in numerous ways. For example, the pharmacological substance Noradrenalin had approximately 350 and 60 different spellings in the Finnish and Swedish data sets, respectively. This problem is part of a more general issue of refer-

ence resolution e.g. when mapping different lexical terms referring to the same concept.

In our quantitative analysis, we have included punctuation characters. In the Swedish data there was a large amount of html-tags and other formatting characters, which has a high impact on the total number of tokens (see Table 2). Moreover, as Finnish is highly inflective, FinCG produces alternative lemmas, hence it is possible to reduce the sparseness of the data by processing the output by choosing only one alternative lemma (see total number of types in Table 2).

To further illustrate the richness of ICU nursing language, the number of unique bigrams (e.g., “*is not*”, “*oxidate well*” and “*night time*” (note: a misspelled compound) are the most common ones for Finnish) and trigrams (e.g., “*oxygenated and ventilated*”, “*and ventilate well*” are among the most common ones for Finnish) were 368,166 (275,205 after FinCG) and 745,407 (356,307 after FinCG) for Finnish patient records. For the Swedish data, the number of unique bigrams was 469,455 (344,127 after GTA) and 1,064,944 (905,539 after GTA). Examples of common Swedish ICU bigrams and trigrams include “*circulation stabile*”, “*during night*”, “*in connection with*”, and “*with good effect*”. Of the content of Finnish nursing narratives, 11% are verbs, 7% nouns and less than 1% pronouns. For Swedish nursing narratives, the respective percentages are 11%, 27% and 2%. One reason for the high numbers for nouns in the Swedish data might be due to the large amount of (obligatory) headings relative to the Finnish data (see Table 3).

To support fluent information flow, language technology is needed to strengthen referential congruence. Much of this richness of vocabulary is explained by abbreviations and personal differences in professional jargon. In particular, abbreviations were common. Based on the analysis of the most common words, abbreviations were relatively established in Swedish data. For the Finnish data, abbreviations were less standard but *RR*, *SR*, *CVP*, *h*, *ad*, *ml*, *ok*, *vas*. [*vasen, left*] and *oik*. [*oikea, right*] were extremely common. Thus, referential congruence can be strengthened by spelling out the most common abbreviations automatically.

Adding topical content headings is another way to support information flow. Topical content headings were mandatory for Swedish data, but no de-

fault headings for Finnish existed. However, the headings for Finnish were established in terms of content. In Table 3, we see that the headings for both languages cover similar topics, which indicates that the clinical information need is similar for professionals in both countries (and languages). Thus, we recommend forming a standardized set of headings from which the user can voluntarily select the ones to be used. This does not exclude adding other headings. Another alternative is to develop language technology for topic segmentation and labeling. We have promising results from this approach (see, e.g., Suominen 2009).

Temporal expressions (e.g., *time*, *evening*, *night*) were often used in both data sets. This poses the question of tense analysis of verbs being unnecessary and the time-related words being enough to imply the needed temporal information. It is also interesting to note that the negations *inte* [not, Swe], *ingen* [none, Swe], *ej* [not, Swe] and *ei* [no/not, Fin] are all among the most common types, which is an important property to take into account in information extraction applications. Furthermore, words regarding the oral cavity, such as *breathing* and *mucus*, as well as relations, such as *daughter*, *son*, *wife*, and *husband* are very common in both data sets.

Inspired by the  $tf \times idf$ -measure from information retrieval, we also analyzed the most common words in terms of a) the number of patients in whose documents the word was used and b) the number of daily nursing narratives in which the word was used. Here, we found, in both data sets, that those words that were used for all patients as well as all daily narratives, were very similar in both data sets, and were related to the most common headings, temporal expressions, negations and monitoring (e.g., *increase*, *continue*, *begin*).

The amount of Protected Health Information (PHI) in form of person names was equal in both of the data sets: 1.5 person names per thousand tokens. This is notable, since this has implications when it comes to integrity issues and reuse of data for research purposes.

FinCG did not recognize 36% of the content of Finnish nursing narratives. However, words marked as unrecognized by FinCG also included punctuation marks. In our previous study (see Suominen 2009 and references therein), we tailored FinCG by extending approximately 35,000 clinical terms. The extension not only substantially

improved the applicability of FinCG to the health domain but also initiated piloting of our language technology components in an authentic healthcare environment in the fall 2008. This led to the release of commercial language technology for Finnish health records (Lingsoft 2010).

Data	Finnish	Swedish
Total number of patients	514	379
Total number of tokens, types (unique tokens) and types after processing	1,227,909 63,328 38,649	1,959,271 - 41,883
Number of tokens per patient:		
Minimum	540	92
Maximum	14,118	36,830
Average	2,389	5,169
Standard deviation	1,635	5,271
Total number of daily documents and shifts	5,915 17,103	4,700 -
Number of tokens per daily document:		
Minimum	0	5
Maximum	915	9,389
Average	208	417
Standard deviation	87	239

Table 2. Comparison of Finnish and Swedish ICU data sets: total amount of text per patient. A daily document, i.e. nursing narrative, contains all text written about a given patient during a calendar day.

Finnish	n ≈	Swedish	n =
Hemodynamics	7,800	Respiratory	11,301
Consciousness	6,900	Circulation	10,630
Relatives	5,700	Elimination	10,041
Diuresis	5,400	Nutrition	8,258
Breathing	4,500	Communication	5,880
Oxygenation	3,600	Event Time	5,681
Other	3,200	Pain	4,732
Excretion	590	Psychosocial	4,682
Hemodialysis	370	Sleep	4,438
Pulse	160	Skin	4,402
Skin	160	Activity	3,794

Table 3. Comparison of Finnish and Swedish ICU data sets: the most common headings. For the Finnish data, where default headings were not given, we approximated the amount of heading by using an automated heuristics followed by manual combination of headings with the same meaning.

For Swedish, GTA handles unknown words differently than FinCG. However, by comparing the ICU words with a Swedish general language corpus (PAROLE, Gellerstam et al. (2000)), we found that 69% of the types are not included in PAROLE, which indicates a need for tailoring GTA (or similar tools for Swedish) with domain-specific ICU terms.

## 4 Conclusions

The purpose of this study was to do content and lexical analysis of nursing narratives written in an ICU. Our findings are that, even though the Finnish and Swedish languages are not linguistically closely related, the way of writing clinical nursing ICU narratives in both countries is very similar. Moreover, the written context made sentences clear for content experts, even though the texts were full of specialized jargon, misspellings, abbreviations, and missing subjects and objects. However, these characteristics make clinical text challenging for language technology. For example coreference resolution as in the case of noradrenalin.

We have also shown that the content characteristics of Finnish and Swedish ICU nursing narratives are very similar. This implies that developing tools for documentation support in ICUs is not country or language dependant in that respect. Developing such tools may improve possibilities for information extraction and text mining, enabling the possibilities to reuse the vast amounts of important practice-based information and evidence captured in clinical narratives. The framework we have introduced here could easily be employed in other studies of clinical texts.

## 6 Future work

In the future, we will use the results of this study in developing language technology for Finnish, Swedish and other Nordic ICU narratives. We will study how to identify abbreviations, misspellings and normalize and correct them, by using various distance measures and concept management techniques. We will also study how to automatically identify important parts of text and highlight them. Furthermore, we are interested in studying text provenance and pragmatics in this particular setting. In addition, we will evaluate the influence of

these technology components in clinical practice. We will also address similarities and differences in clinical text written by various professional groups or at other hospital wards and health care units. Finally, we are eager to seek possibilities to incorporate laymen's information needs and their interaction with health care providers into our study.

## Acknowledgments

We would like to thank Nordforsk and the Nordic Council of Ministers for the funding of our research network HEXAnord – HHealth teXt Analysis network in the Nordic and Baltic countries and NICTA, funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. We would also like to thank the Department of Information Technology and TUCS, University of Turku, Finland.

## References

- Lars Borin, Natalia Grabar, Catalina Hallett, Davis Hardcastle, Maria Toporowska Gronostaj, Dimitrios Kokkinakis, Sandra Williams and Alistair Willis. 2007. Empowering the patient with language technology. SemanticMining NoE 507505: *Deliverable D27.2*. <<http://gup.ub.gu.se/gup/record/index.xhtml?pu bid=53590>>
- Grace Yuet-Chee Chung. 2009. Towards identifying intervention arms in randomized controlled trials: extracting coordinating constructions. *Journal of Biomedical Informatics*. 42(5):790–800
- Hercules Dalianis, Martin Hassel and Sumithra Velupilai. 2009. The Stockholm EPR Corpus - Characteristics and Some Initial Findings. *Proceedings of ISHIMR 2009*, Evaluation and implementation of e-health and health information initiatives: international perspectives. 14th International Symposium for Health Information Management Research, Kalmars, Sweden, 14-16 October, 2009, pp 243-249, pdf. Awarded best paper.
- Anna Ehrenberg, Margareta Ehnfors and Ingrid Thorell-Ekstrand. 1996. Nursing documentation in patient records: experience of the use of the VIPS model. *Journal of Advanced Nursing* 24, 853–867.
- Martin Gellerstam, Yvonne Cederholm, and Torgny Rasmark. The bank of Swedish. In: *Proceedings of LREC 2000 -- The 2nd International Conference on Language Resources and Evaluation*, pages 329–333, Athens, Greece.
- Udo Hahn and Joachim Wermter. 2004. High-performance tagging on medical texts. *Proceedings of the 20th international conference on Computational Linguistics*. Geneva, Switzerland.

- Henk Harkema, Dowling JN, Thornblade T, Chapman WW. 2009. ConText: an algorithm for determining negation, experienter, and temporal status from clinical reports. *Journal of Biomedical Informatics* 2009;42(5):839–51.
- Ragnhild Hellesø. 2005. Information handling in the nursing discharge notes, *Journal of Clinical Nursing*, Volume 15 Issue 1, 11 - 21. Blackwell publishing
- Gunlög Josefsson. 1999. Få feber eller tempa? Några tankar om agentivitet i medicinskt fackspråk, *Alla tiders språk: en vänskrift till Gertrud Pettersson*. Pages 127. Institutionen för nordiska språk. Lund. (In Swedish)
- Hyeoneui Kim, Sergey Goryachev, Craciela Rosemblat, Allen Browne, Alla Keselman and Qing Zeng-Treitler. 2007. Beyond surface characteristics: a new health text-specific readability measurement. *AMIA Annual Symp.* 11:418-22.
- Ola Knutsson, Johnny Bigert, and Vigg Kann. 2003. A robust shallow parser for Swedish. In *Proceedings 14th Nordic Conf. on Comp. Ling. NODALIDA*.
- Sirkka Lauri and Sanna Salanterä. 2002. Developing an instrument to measure and describe clinical decision making in different nursing fields. *Journal of Professional Nursing*. Mar-Apr;18(2), 93-100.
- Lingsoft. 2010, Lingsoft Oy, <http://www.lingsoft.fi/>
- Clement J. McDonald. 1997. The Barriers to Electronic Medical Record Systems and How to Overcome Them. *JAMIA*. 1997;4:213–221.
- Stéphane M. Meystre, Guergana K. Savova, Karin C. Kipper-Schuler and John E. Hurdle. 2008. Extracting Information from Textual Documents in the Electronic Health Record: a Review of Recent Research. *Yearbook Med Inform.* 2008:128-44.
- Raymond L. Ownby 2005. Influence of Vocabulary and Sentence Complexity and Passive Voice on the Readability of Consumer-Oriented Mental Health Information on the Internet. *AMIA Annual Symposium Proceedings*. 2005: 585–588.
- Serguei V. S. Pakhomov, Anni Coden and Christopher G. Chute. 2006. Developing a corpus of clinical notes manually annotated for part-of-speech. *International Journal of Medical Informatics*. 2006 Jun;75(6):418-29. Epub 2005 Sep 19.
- Patientdatalagen (2008:355) *Svensk författningssamling*, Socialdepartementet, 2008, Stockholm. (In Swedish)
- Hanna Suominen. 2009. Machine Learning and Clinical Text: Supporting Health Information Flow. *TUCS Dissertations* No 125, Turku Centre for Computer Science, 2009, Turku, Finland.
- Hanna Suominen, Heljä Lundgrén-Laine, Sanna Salanterä, Helena Karsten, and Tapio Salakoski. 2009. Information flow in intensive care narratives. In Chen J, Chen C, Ely J, Hakkani-Tr D, He J, Hsu H.-H, Liao L, Liu C, Pop M, Ranganathan S, Reddy C.K, Ruan J, Song Y, Tseng V.S, Ungar L, Wu D, Wu Z, Xu K, Yu H, Zelikovsky A, editors. *Proceedings IEEE International Conference on Bioinformatics and Biomedicine Workshops*, BIBM 2009, pages 325–330. Institute of Electrical and Electronics Engineers, Los Alamitos, California, USA.
- Kaarina Tanttu and Helena Ikonen. 2007. Nationally standardized electronic nursing documentation in Finland by the year 2007. *Stud Health Technol Inform.*122:540-1.
- Asta Thoroddsen, Kaija Saranto, Anna Ehrenberg, Walter Sermeus. 2009. Models, standards and structures of nursing documentation in European countries. *Stud Health Technol Inform.*146:327-31.
- Katrin Tomanek, Joachim Wermter and Udo Hahn. 2007. A Reappraisal of sentence and token splitting for life sciences documents. *Stud Health Technol Inform.* 129 (Pt 1):524-8.
- Webster's 2010. Webster's New World Medical Dictionary. <http://www.medterms.com/script/main/hp.asp>, last visited February 2, 2010.