

Medical diagnosis lost in translation – Analysis of uncertainty and negation expressions in English and Swedish clinical texts

Danielle L Mowery
University of Pittsburgh
200 Meyran Ave
Pittsburgh, PA 15260
dlm31@pitt.edu

Sumithra Velupillai
Stockholm University
164 40 Kista
Stockholm, Sweden
sumithra@dsv.su.se

Wendy W Chapman
University of California San Diego
10100 Hopkins Dr
La Jolla, CA 92093
wwchapman@ucsd.edu

Abstract

In the English clinical and biomedical text domains, negation and certainty usage are two well-studied phenomena. However, few studies have made an in-depth characterization of uncertainties expressed in a clinical setting, and compared this between different annotation efforts. This preliminary, qualitative study attempts to 1) create a clinical uncertainty and negation taxonomy, 2) develop a translation map to convert annotation labels from an English schema into a Swedish schema, and 3) characterize and compare two data sets using this taxonomy. We define a clinical uncertainty and negation taxonomy and a translation map for converting annotation labels between two schemas and report observed similarities and differences between the two data sets.

1 Introduction and Background

Medical natural language processing techniques are potentially useful for extracting information contained in clinical texts, such as emergency department reports (Meystre et al., 2008). One important aspect to take into account when developing accurate information extraction tools is the ability to distinguish negated, affirmed, and uncertain information (Chu et al., 2006). Several research studies have targeted this problem and created annotation schemas and manually annotated reference standards for uncertainty and negation occurrence in news documents (Saurí and Pustejovsky (2009), Wiebe et al. (2001), Rubin et al. (2006)), biomedical research articles (Wilbur et al. (2006), Vincze et al.

(2008)), and clinical narratives (Uzuner et al. (2011) and Uzuner et al. (2009)). There are encoding tools developed for automatic identification of uncertainty and negation in English, such as ConText (Harkema et al., 2009), which relies on heuristics and keyword lists, and MITRE’s CARAFE (Clark et al., 2011), which combines heuristic and statistical techniques.

However, most relevant annotation schemas, reference standards, and encoding tools are built for English documents. For smaller languages, such as Swedish, resources are scarce.

We present a pilot, qualitative study to compare two different annotation schemas and subsequent annotated corpora for uncertainty modeling of disorder mentions, e.g., signs, symptoms, and diseases, in clinical texts, for two different languages: English and Swedish. We compare these annotation schemas and their instantiation in the two languages in an attempt to gain a deeper understanding of how uncertainty and negation are expressed in different clinical texts with an emphasis on creating a portable uncertainty and negation application that generalizes among clinical texts of different languages.

This pilot study is motivated for at least two reasons. First, little attention has been given to mapping, characterizing, or comparing annotation schemas built for different languages or to characterizing different types of uncertainty expressions and the intention underlying those expressions. Such knowledge is needed for building information extraction tools that can accurately identify or track differential diagnoses over time, particularly when medical reasoning can be laden with uncertainty about a disorder’s existence or change over time.

Second, building new resources for small languages is time consuming. Utilizing existing tools and techniques already developed for one language, such as English, could be an efficient way of developing new useful tools for other less exploited languages, such as Swedish.

Our overall goal is to move towards improving automatic information extraction from clinical texts by leveraging language differences and similarities. In order to address this issue, our aims in this study are to 1) create a taxonomy for deepened characterization of how uncertainty and negation is expressed in clinical texts, 2) compare two existing uncertainty and negation annotation schemas from this perspective, and 3) compare differences and similarities in expressions of uncertainty and negation between two languages: English and Swedish.

2 Methods

In this pilot, qualitative comparison study, we used grounded theory (Strauss and Corbin, 1990) to inductively identify themes that characterize clinical uncertainty and negation expressed in both English (University of Pittsburgh Medical Center) and Swedish (Karolinska University Hospital) research data sets derived from emergency department reports.

2.1 Uncertainty/negation annotation schemas

Two independently developed annotation schemas were used to annotate disorder mentions in the clinical texts: a schema developed for English reports (Mowery et al. (2012)) and one for Swedish (Velupillai et al. (2011)). Each disorder mention was pre-annotated and constituted the input to a separate set of annotators, who assigned values to a set of attributes defined in the schema. For instance, in the sentence “Patient with possible pneumonia.”, annotators for the English data set assigned **values** to four attributes for the instance of pneumonia:

- Existence(**yes**, no): whether the disorder was ever present
- AspectualPhase(initiation, continuation, culmination, **unmarked**): the stage of the disorder in its progression
- Certainty(low, **moderate**, high, unmarked): amount of certainty expressed about whether the disorder exists
- MentalState(**yes**, no): whether an outward thought or feeling about the disorder’s existence is mentioned

In the Swedish schema, annotators assigned **values** to two attributes:

- Polarity(**positive**, negative): whether a disorder mention is in the positive or negative polarity, i.e., affirmed (positive) or negated (negative)
- Certainty(**possibly**, probably, certainly): gradation of certainty for a disorder mention, to be assigned with a polarity value.

2.2 Data Sets

The English data set included 30 de-identified, full-length emergency department reports annotated with 283 disorders related to influenza-like illnesses by a board-certified infectious disease physician. Each disorder was annotated with four attributes – existence, aspectual phase, certainty and mental state – by two independent annotators (including DM) who came to consensus after reviewing disagreements.

The Swedish data set included 1,297 assessment sections from emergency department reports annotated with approx. 2,000 disorders, automatically marked from a manually created list of approximately 300 unique disorders by two physicians. The two physicians annotated each disorder mention with attributes of polarity and certainty. A random subset of approx. 200 annotated disorder mentions from the data set were used for this qualitative study.

2.3 Study Process

In order to better understand how physicians describe uncertainty of the presence or absence of a disorder, we evaluated the annotations from the two data sets as follows: 1) created a clinical uncertainty and negation taxonomy, 2) developed a translation map for mapping attributes and values from the English schema into the Swedish schema, and 3) characterized and compared both data sets and languages using the taxonomy.

To create the uncertainty and negation taxonomy, we conducted a literature review of recent annotation schemas (e.g. Vincze et al. (2008)), assignment applications (e.g. Uzun et al. (2011), Harkema et al. (2009), Clark et al. (2011), Chapman et al. (2011)), and observational studies (Lingard et al., 2003) about uncertainty or negation in the clinical domain. From our review, we created a clinical taxonomy describing notable characteristics of uncertainty and negation, which were added to and refined using grounded theory, by inspecting the disorder annotations in our data sets and documenting

emerging themes consistent with issues found from the literature review. For instance, one characteristic of negation annotations found in the literature and in our data sets is the existence of a lexical cue indicating that a disorder is negated, and the lexical cue can occur before, within, or after the disorder mention. The characteristics included in the taxonomy represent *features* describing the attributes of uncertainty and negation in the data sets (see Section 3.1).

To develop the translation map between certainty and negation values from each annotation schema, authors DM and SV jointly reviewed each annotated disorder mention from the English data set and assigned a Swedish polarity and certainty label, then devised a map from the English schema into the Swedish schema.

To characterize and compare manifestations of uncertainty and negation using annotations from the two data sets, DM and SV annotated each disorder mention in both data sets with the features in the clinical uncertainty and negation taxonomy. In the English data set, each disorder was annotated by DM and adjudicated by SV. In the Swedish data set, each disorder was annotated by SV then translated into English for adjudication by DM.

3 Results

3.1 Clinical Uncertainty and Negation Taxonomy

We developed a clinical uncertainty and negation taxonomy to characterize the linguistic manifestations of uncertainty and negation in clinical text (Figure 1). We found three high-level features in the literature and in our data sets: *position of lexical cue* (i.e., position of the lexical expression indicating uncertainty or negation with respect to the disorder), *opinion source* (i.e. person believing there is absence, presence, or uncertainty), and *evidence evaluation* (i.e., reason for the uncertainty or negation belief).

Position of lexical cue demonstrated itself in the data sets in three non-mutually exclusive ways:

- *pre-disorder* (lexical cue precedes the disorder) “Patient denies chest pain.”
- *intra-disorder* (lexical cue occurs within the name of the disorder) “x-ray...possibly be indicative of pneumonia.”

- *post-disorder* (lexical cue occurs after the disorder) “abdominal cramping..is unlikely.”

Opinion source exhibited the following values:

- *dictating physician* (dictating physician alone expressed presence, absence, or uncertainty regarding the disorder) “I suspect bacterial pneumonia.”
- *dictating physician with consultation* (dictating physician explicitly includes other clinical professional in statement) “Discussing with Dr. **NAME**, pneumonia can not be excluded.”
- *other clinical care providers* (other clinical team members explicitly stated as expressing presence, absence or uncertainty regarding the disorder) “per patient’s primary doctor, pneumonia is suspected.”
- *patient* (patient expressed presence, absence, or uncertainty regarding the disorder) “Pt doesn’t think she has pneumonia.”
- *unknown* (ambiguous who is expressing presence, absence, or uncertainty regarding the disorder) “there was a short episode of coughing.”

Evidence evaluation includes a modified subset of values found in the model of uncertainty proposed by Lingard et al. (2003) to connote perceived reasons for the provider uncertainty (and negation) about the disorder mention as used in our data sets.

- *limits of evidence* (data limitations for hypothesis testing), *one diagnosis*
 - *evidence contradicts* (data contradicts expected hypothesis), “Blood test normal, but we still think Lyme disease.”
 - *evidence needed* (evidence unavailable to test hypothesis) “Waiting for x-ray results to determine if it’s a femur fracture.”
 - *evidence not convincing, but diagnosis asserted* (data doesn’t fully support proposed hypothesis), “Slightly elevated levels of WBCs suggests infection.”
- *limits of evidence, more than one diagnosis*
 - *differential diagnoses enumerated* (competing diagnoses reasoned), “bacterial infection vs. viral infection.”
- *limits in source of evidence* (untrusted evidence)
 - *non-clinical source* (from non-provider source), “Pt can’t remember if she was diagnosed with COPD.”
 - *clinical source* (from provider source), “I do not agree with Dr. X’s diagnosis of meningitis.”
 - *test source* (from test e.g., poor quality), “We cannot determine from the x-ray if the mass is fluid or a tumor.”
- *limitless possibilities* (large number of likely diagnoses so diagnosis defaulted to most likely), “This is probably an infection of some sort.”
- *other* (no evidence limitation)
 - *asserting a diagnosis or disorder as affirmed* (positive case), “Confirms nausea.”
 - *asserting a diagnosis or disorder as negated* (negative case), “No vomiting.”

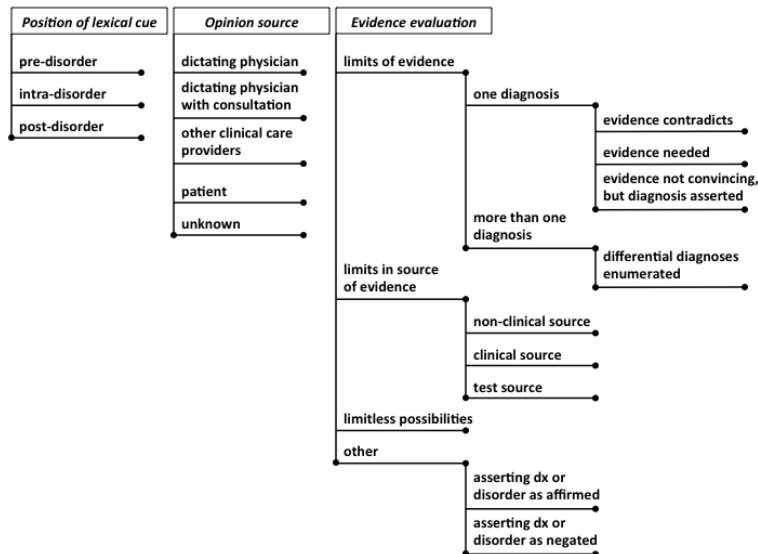


Figure 1: Uncertainty and negation taxonomy with features – *Position of lexical cue*, *Opinion source* and *Evidence evaluation* – with corresponding values (nested lines and sub-lines).

3.2 Translation Map

In order to compare annotations between the data sets, we developed a mapping procedure for converting the four annotated attribute values from the English schema into the two annotated attribute values from the Swedish schema. This mapping procedure uses two normalization steps, negation and certainty (see Figure 2).

Using Figure 2, we explain the mapping procedure to convert English annotations into Swedish annotations. Our steps and rules are applied with precedence, top down and left to right. For “I have no suspicion for bacterial infection for this patient”, English annotations are Existence(no) AND AspectualPhase(null) AND Certainty(high) AND MentalState(yes), and Swedish annotations are Polarity(negative) AND Certainty(probably). The mapping procedure applies two normalization steps, negation and uncertainty, with the following rules.

The **first step** is negation normalization to *convert Existence and Aspectual Phase into Polarity annotations*. In this example, **Existence(no)** → **Polarity(negative)**.

The **second step** is certainty normalization with up to two sub steps. For Certainty mapping, in summary, *map English NOT Certainty(unmarked) to Swedish Certainty level*, e.g., **Certainty(high)** → **Certainty(probably)**. For MentalState

mapping, *if English Certainty(unmarked) AND MentalState(yes), map to either Swedish Certainty(probably) OR Certainty(possibly) using your best judgment; otherwise, map to Certainty(certainly)*. For our example sentence, Certainty mapping was sufficient to map from the English to the Swedish Certainty levels.

We found that these two schemas were mappable. Despite the binary mapping splits from English Certainty(Moderate) → Swedish Certainty(possibly) OR Certainty(probably) and judgment calls necessary for MentalState mapping, few annotations were not easily mapped.

3.3 Characterization of English and Swedish Data sets

In this study, we characterized our data sets according to a clinical uncertainty and negation taxonomy comprised of three concepts – *position of lexical cue*, *opinion source*, and *evidence evaluation*.

3.3.1 Position of lexical cue

In Table 1, we show examples of phrases from each data set representing the Polarity and Certainty levels in the taxonomy. In our data set, we did not explicitly annotate markers for the highest certainty levels in the positive polarity, such as “definitely has”. We did not encounter any of these cases in the

"I have *no suspicion* for bacterial infection for this patient."

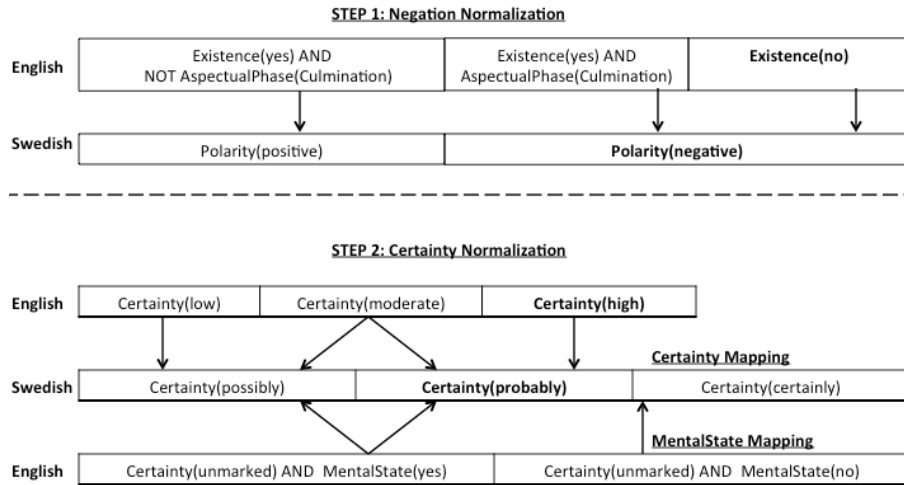


Figure 2: Map between values for attributes in Swedish and English schemas. Bolded rules indicate the rules used to assign values to the example sentence (English sentence on top).

data set. We observed that most uncertainty expressions precede a disorder mention. Few expressions both precede and follow the disorder mention, or within the disorder mention itself. We observed that most expressions of uncertainty are conveyed using positive polarity gradations such as “probably” and “possibly”, for example “likely”, “appears to have”, “signs of”. Lexical cues of low levels of certainty in the negative polarity were rare.

3.3.2 Opinion source

In Table 2, we report examples of the various individuals – *dictating physician, dictating physician with consultation, other clinical care providers, patient, unknown* – that are the source of the belief state for uncertainty about a disorder. We observed explicit judgments or mental postulations e.g., “I judge” or implied speculations in which the physician was not the subject and passive expressions were used e.g., “patient appears to have”. In cases of dictating physician with consultation, the physician speculated about the disorder using references to other providers consulted to strengthen the assessment e.g., “Discussing with Dr..”. In cases of other clinical care providers, there was no ownership on the part of the dictating physician, but of other members of the clinical care team e.g., “Con-

sulting Attending (Infection) thinks...”. In cases for patient, the patient is conveying statements of confusion with respect to self-diagnosing e.g., “Pat. reports that she finds it difficult to discern...”. We observed no expressions of uncertainty owned by the patient in the English data set or by a relative in the Swedish data set. In the unknown case, it is unclear from the context of the report whether the speculation is on the part of the physician to believe the symptom reported or the relative unsure about reporting the symptoms e.g., “there was apparently”.

3.3.3 Evidence evaluation

Below we list examples of the different reasons for uncertainties that were identified. Not all types were observed in both corpora (Not observed).

limits of evidence, one diagnosis

- evidence contradicts – **English**: “Likely upper GI bleed with elevated bun, but normal h and h.”; **Swedish**: “Kon-sulterar infektionsjour som anser viros vara osannolikt med tanke på normalt leverstatus. (Consulting Attending (infection) who thinks that virosis is improbable given normal liver status.)”
- evidence needed – **English**: “chest x-ray was ordered to rule out TB.”; **Swedish**: “Diskuterar med RAH-jour; vi börjar utredning med CT-skalle med kontrast på misstanke om metastaser och någon form av epileptiskt anfall (Discussion with Attendant [CLINIC]; we start inves-

	Certainly		Probably		Possibly	
	English	Swedish	English	Swedish	English	Swedish
Positive	<default>	<default>	<ul style="list-style-type: none"> likely suspect thought it was felt appears to have there was apparently most likely 	<ul style="list-style-type: none"> förmodligen, troligen(probably) troligtvis, troligen (probably/likely) [most] sannolikt ([most] probable) tecken på (signs of) oklar (unclear) 	<ul style="list-style-type: none"> this is not unlikely rule out differential includes possible could possible be indicative of possible I think this is probably probable 	<ul style="list-style-type: none"> möjlig[en] [twis], (possibly) eventuell, ev, möjlig (possible) misstanke [på] (suspicion [for]) skulle kunna vara (could be) kan [ej]inte uteslutas (cannot be ruled out)
Negative	<ul style="list-style-type: none"> denies; denies any recent no, not has never had; has not had no further episodes of no prior episode stopped resolution of has abated resolved; resolves; had resolved clear; free unremarkable for evidence of 	<ul style="list-style-type: none"> ingen misstanke [om] för] (no suspicion for) ing[en] [a] (no) inga hållpunkter för (no indication of) utesluter (rule out) 	<ul style="list-style-type: none"> no suspicion for 	<ul style="list-style-type: none"> ingen stark [klinisk]misstanke [om] (no strong clinical suspicion for) ej visar tecken på (does not show signs for) 	<ul style="list-style-type: none"> am not convinced 	

Table 1: Common lexical cues and their relative position to the disorder mention: *Pre-disorder*: uncertainty marker before disorder, *Intra-disorder*: uncertainty marker inside disorder, *Post-disorder*: uncertainty marker after disorder, †= schema compatibility/neutral case.

Opinion source	English	Swedish
Dictating physician	<ul style="list-style-type: none"> it seems to me there is some active upper GI bleeding. 	<ul style="list-style-type: none"> ...jag bedömer dock inte att besvären tyder på någon ungemboll... (...however I do not judge that the problems indicate any pulmonary embolism...)
Dictating physician with consultation	<ul style="list-style-type: none"> ...evaluation by myself and Dr. **NAME. It was felt that the patient was suffering from viral pharyngitis. 	<ul style="list-style-type: none"> Diskuterar med RAH-jour; vi bärjar utredning med CT-skalle med kontrast på misstanke om metastaser och någon form av epileptiskt anfall (Discussion with oncologist on call; we start investigation with CT-brain with contrast on suspicion of metastases and some kind of epileptic seizure.)
Other clinical care provider	<ul style="list-style-type: none"> ...per his primary care doctor...they thought it was a viral syndrome of unknown type. 	<ul style="list-style-type: none"> Konsulterar infektionsjour som anser virus vara osannolikt med tanke på normalt leverstatus. (Consulting Attending (infection) who thinks that virosis is improbable given normal liver status.)
Patient	--	<ul style="list-style-type: none"> ... hon själv tror att det kan röra sig om panikångestattacker... (she herself thinks that it can be panick attacks...)
Unknown	<ul style="list-style-type: none"> there was apparently a short episode of apnea... 	--

Table 2: Opinion source of uncertainty or negation types with English and Swedish examples.

tigation with CT-brain with contrast on suspicion for metastasis and some form of epileptic seizure.)”

- evidence not convincing, but diagnosis asserted – **English**: Not observed; **Swedish**: “Förmodligen en viros eftersom man kan se en viss lymfocytopeni i diff (Probably a virosis since there is some lymphocyte in blood cell count.)”

limits of evidence, more than one diagnosis

- differential diagnoses enumerated – **English**: “questionable right-sided increased density on the right side of the chest x-ray that could possibly be indicative of a pneumonia versus increased pulmonary vasculature”; **Swedish**: “Förefaller neurologiskt, blödning? Infarkt? (Appears neurological, bleeding? Infarction?)”

limits in source of evidence

- non-clinical source – **English**: “I am not convinced that he is perfectly clear on his situation..”; **Swedish**: “Pat

uppper att hon har svårt att skilja på panikångest och andra symtom. (Pat. reports that she finds it difficult to discern panick disorder from other symptoms...)”

- clinical source – **English**: “there was no definite diagnosis and they thought it was a viral syndrome of unknown type.”; **Swedish**: Not observed
- test source – **English**: “..confusion was possible related a TIA without much facial droop appreciated on my physical exam”; **Swedish**: “Ter sig mest sannolikt som reumatoid artrit både klinisk och lab-mässigt (Seems like it most probably is rheumatoid arthritis both clinically and lab-wise.)”

limitless possibilities – **English**: “I think this is probably a viral problem.”; **Swedish**: “Pat bedömes ha en förkylning, troligen virusinfektion. (Patient is evaluated as having a cold, probably a virus infection.)”

other

- *asserting dx or disorder as affirmed* – **English:** “*I suspect that colon cancer is both the cause of the patient’s bleeding.*”; **Swedish:** Not observed
- *asserting dx or disorder as negated* – **English:** “*...her fever has abated.*”; **Swedish:** Not observed

In many cases, the local context was sufficient for understanding the evidential origins for uncertainty. When a single disorder was mentioned, uncertainty was due to data insufficient to make a definitive diagnosis because it contradicted a hypothesis, was unavailable, or was not convincing. For instance, data was to be ordered and the opportunity to interpret it had not presented itself, such as “..was ordered to rule out TB” or “..start investigation with CT-brain with contrast..”. In few cases, more than one diagnosis was being enumerated due to a limitation in the evidence or data gathered e.g., “Appears neurological, bleeding? Infarction?”. We observed cases in which the source of the evidence produced uncertainty including both non-clinical and clinical sources (care providers consulted and tests produced). In cases of limitless possibilities, the physician resorted to a common, default diagnosis e.g., “probably a virus infection”. Limitations of evidence from a clinical source were not found in the Swedish data set and few were found in the English data set. We expect that more examples of this category would be found in e.g. radiology reports in which the quality of the image is a critical factor in making an interpretation.

4 Discussion and Conclusion

From the resulting clinical taxonomy and characterization, we observe some general differences and similarities between the two data sets and languages. The Swedish assessment entries are more verbose compared to the English medical records in terms of a more detailed account of the uncertainty and what is being done by whom to derive a diagnosis from a disorder mention. This might reflect cultural differences in how documentation is both produced and used. Differential diagnoses are often listed with question marks (“?”) in the Swedish set, e.g., “Disorder 1? Disorder 2? Disorder 3?”, whereas in the English data set enumerations are either listed or competing, e.g., “disorder 1 vs. disorder 2”. Despite these differences, there are many similarities

between the two data sets.

Mapping observations from the English schema into the Swedish schema was not complicated despite the difference in the modeled attributes. In most cases, we determined that designating attribute-value rules for negation and certainty normalization steps was sufficient to accurately map observations between the language schemas without changing an observation’s semantics. This finding suggests that simple heuristics can be used to translate annotations made from English trained tools into the Swedish schema values.

The majority of the lexical markers are positioned in both languages, and the majority of these markers are similar across the two languages, e.g., “likely”, “possible”, “suspicion for”. However, inflections and variants are more common in Swedish, and the language allows for free word order, this relation needs to be studied further. The default case, i.e. affirmed, or certainly positive, was rarely expressed through lexical markers.

When it comes to the opinion source of an uncertainty or negation, we observed a pattern in the use of passive voice, e.g. “it was felt”, indicating avoidance to commitment in a statement. Accurate extraction of the opinion source of an expression has important implications for a system that, for instance, tracks the reasoning about a patient case over time by source. This has been recognized and incorporated in other annotation efforts, for example for news documents (Saurí and Pustejovsky, 2009). In the English data set, no cases of self-diagnosing are found, i.e. a patient owning the expressed uncertainty. In both data sets, an implicit dictating physician source is most common, i.e. there is no explicit use of pronouns indicating the opinion holder. In most cases it is clear that it is the writer’s (i.e. the dictating physician’s) opinion that is expressed, but in some cases, a larger context is needed for this knowledge to be resolved.

Reviewing the evidential origins or reason for expressed uncertainty, for both the Swedish and English data sets, the category “limits of evidence” is most common. This reflects a clinical reality, where many disorders require test results, radiology findings and other similar procedures before ascertaining a diagnosis. Although most cases of uncertainty are manifested and strengthened through a lexical

marker, there are also instances where the uncertainty is evident without such explicit markers, e.g. the ordering of a test may in itself indicate uncertainty.

4.1 Limitations

There are several limitations of this study. The Swedish data set only contains parts of the medical record and the English data set is very small. In the creation of the taxonomy and characteristics, we have not focused on studying uncertainty *levels*, i.e. distinctions between “possibly” and “probably”. The values of our taxonomy are preliminary and may change as we develop the size of our data set. Additionally, we only studied emergency department reports. We need to study other report types to evaluate the generalizability of the taxonomy.

The two compared languages both origin from the same language family (Germanic), which limits generalizability for other languages. Furthermore, the definitions of disorders in the two sets differ to some extent, i.e., English disorders are related to specific influenza-like illnesses and Swedish to more general disorders found in emergency departments.

4.2 Comparison to related work

Annotation schemas and reference standards for uncertainty and negation have been created from different perspectives, for different levels and purposes. The BioScope Corpus, for instance, contains sentence-level uncertainty annotations with token-level annotations for speculation and negation cues, along with their linguistic scope (Vincze et al., 2008). In Wilbur et al. (2006), five qualitative dimensions for characterizing biomedical articles are defined, including levels of certainty. In the 2010 i2b2/VA Challenge on concepts, assertions and relations in clinical texts, medical problem concepts were annotated. The assertion task included six annotation classes (present, absent, possible, hypothetical, conditional, not associated with the patient), to be assigned to each medical problem concept (Uzuner et al., 2011). Vincze et al. (2011) present a quantitative comparison of the intersection of two English corpora annotated for negation and speculation (BioScope and Genia Event) from two different perspectives (linguistic and event-oriented).

We extend these schemas by characterizing the underlying meaning and distinctions evident by the linguistic expressions used to indicate uncertainty and negation in the clinical domain and by exploring the relationship between uncertainty and negation, through an analysis and comparison of two different annotation schemas. However, this study is not a proposal for mapping to these schemas or others.

From an application perspective, uncertainty and negation handling have been included in rule-based systems such as NegEx and ConText, applied on disorder mentions. In Chapman et al. (2011), a generalized version of ConText is presented, with uncertainty values (probably, definitely) linked to either a positive or negative assertion, with an added indeterminate value. A previous study has shown promising results for adapting NegEx to Swedish (Skeppstedt, 2011), indicating that further extensions and adaptations between the two languages for e.g. uncertainty modeling should be viable. Machine-learning based approaches outperform rule-based for assertion classification according to results presented in Uzuner et al. (2009). A machine-learning approach was also used in the top performing system in the 2010 i2b2/VA Challenge assertion task (de Bruijn et al., 2011).

4.3 Implications and future work

With uncertainty lexicons for both Swedish and English, we hypothesize that we will be able to extend ConText to handle uncertainties in English as well as in Swedish. This enables both improvements over the existing system and the possibilities of further comparing system performances between languages. We will also experiment with machine-learning approaches to detect and annotate uncertainty and negation. We plan to extend both data sets, the English data set using semi-automatically translated disorders marked in the Swedish data set to encode new disorder mentions, and the Swedish data set by extracting the full medical records, thus creating a larger set for comparison. We will extend the taxonomy as needed e.g., syntactic and semantic patterns, and investigate how to integrate the clinical taxonomy to inform ConText by providing more granular descriptions of the motivation behind the uncertainty, thus bringing us closer to natural language understanding.

Acknowledgments

For the English and Swedish data sets, we obtained approval from the University of Pittsburgh IRB and the Regional Ethical Review Board in Stockholm (Etikprövningsnämnden i Stockholm). The study is part of the Interlock project, funded by the Stockholm University Academic Initiative and partially funded by NLM Fellowship 5T15LM007059. Lexicons and probabilities will be made available and updated on the iDASH NLP ecosystem under Resources: <http://idash.ucsd.edu/nlp/natural-language-processing-nlp-ecosystem>.

References

- B. E. Chapman, S. Lee, H. Peter Kang, and W. W. Chapman. 2011. Document-level Classification of CT Pulmonary Angiography Reports Based on an Extension of the ConText Algorithm. *Journal of Biomedical Informatics*, 44:728–737.
- D. Chu, J.N. Dowling, and WW Chapman. 2006. Evaluating the Effectiveness of Four Contextual Features in Classifying Annotated Clinical Conditions in Emergency Department Reports. In *AMIA Annu Symp Proc*, pages 141–145.
- C. Clark, J. Aberdeen, M. Coarr, D. Tresner-Kirsh, B. Wellner, A. Yeh, and L. Hirschman. 2011. MITRE system for Clinical Assertion Status Classification. *J Am Med Inform Assoc*, 11(18):563–567.
- B. de Bruijn, C. Cherry, S. Kiritchenko, J. Martin, and X. Zhu. 2011. Machine-learned Solutions for Three Stages of Clinical Information Extraction: The State of the Art at i2b2 2010. *Journal of the American Medical Informatics Association*, 18:557–562.
- H. Harkema, J. N. Dowling, T. Thornblade, and W. W. Chapman. 2009. ConText: An Algorithm for Determining Negation, Experiencer, and Temporal Status from Clinical Reports. *Journal of Biomedical Informatics*, 42:839–851.
- L. Lingard, K. Garwood, C. F. Schryer, and M. M. Spafford. 2003. A Certain Art of Uncertainty: Case Presentation and the Development of Professional Identity. *Social science medicine*, 56(3):603–616.
- S. M. Meystre, G. K. Savova, K. C. Kipper-Schuler, and John E. Hurdle. 2008. Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. *IMIA Yearbook of Medical Informatics 2008. 47 Suppl 1:138-154*.
- D. Mowery, P. Jordan, J.M. Wiebe, H. Harkema, and W.W. Chapman. 2012. Semantic Annotation of Clinical Text: A Pilot Study. Unpublished.
- V. L. Rubin, E. D. Liddy, and N. Kando. 2006. Certainty Identification in Texts: Categorization Model and Manual Tagging Results. In *Computing Affect and Attitude in Text: Theory and Applications*. Springer.
- R. Saurí and J. Pustejovsky. 2009. FactBank: A Corpus Annotated with Event Factuality. *Language Resources and Evaluation*, 43(3):227–268–268, September.
- M. Skeppstedt. 2011. Negation Detection in Swedish Clinical Text: An Adaptation of NegEx to Swedish. *Journal of Biomedical Semantics*, 2(Suppl. 3):S3.
- A. L. Strauss and J. Corbin. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage.
- Ö. Uzuner, X. Zhang, and T. Sibanda. 2009. Machine Learning and Rule-based Approaches to Assertion Classification. *Journal of the American Medical Informatics Association*, 16(1):109–115.
- Ö. Uzuner, B. R. South, S. Shen, and S. L. DuVall. 2011. 2010 i2b2/VA Challenge on Concepts, Assertions, and Relations in Clinical Text. *JAMIA*, 18(5):552–556.
- S. Velupillai, H. Dalianis, and M. Kvist. 2011. Factualty Levels of Diagnoses in Swedish Clinical Text. In A. Moen, S. K. Andersen, J. Aarts, and P. Hurlen, editors, *Proc. XXIII International Conference of the European Federation for Medical Informatics (User Centred Networked Health Care)*, pages 559 – 563, Oslo, August. IOS Press.
- V. Vincze, G. Szarvas, R. Farkas, G. Móra, and J. Csirik. 2008. The BioScope Corpus: Biomedical Texts Annotated for Uncertainty, Negation and Their Scopes. *BMC Bioinformatics*, 9(S-11).
- V. Vincze, G. Szarvas, G. M’ora, T. Ohta, and R. Farkas. 2011. Linguistic Scope-based and Biological Event-based Speculation and Negation Annotations in the BioScope and Genia Event Corpora. *Journal of Biomedical Semantics*, 2(Suppl. 5):S8.
- J. Wiebe, R. Bruce, M. Bell, M. Martin, and T. Wilson. 2001. A Corpus Study of Evaluative and Speculative Language. In *Proceedings of the Second SIGDIAL Workshop on Discourse and Dialogue - Volume 16, SIGDIAL ’01*, pages 1–10, Stroudsburg, PA, USA. Association for Computational Linguistics.
- J. W. Wilbur, A. Rzhetsky, and H. Shatkay. 2006. New Directions in Biomedical Text Annotation: Definitions, Guidelines and Corpus Construction. *BMC Bioinformatics*, 7:356+, July.