

Medical text simplification using synonym replacement: Adapting assessment of word difficulty to a compounding language

Emil Abrahamsson¹ Timothy Forni¹ Maria Skeppstedt¹ Maria Kvist^{1,2}

¹Department of Computer and Systems Sciences (DSV)
Stockholm University, Sweden

{emab6827, tifo6794, mariask}@dsv.su.se

²Department of Learning, Informatics, Management and Ethics (LIME)
Karolinska Institutet, Sweden

maria.kvist@karolinska.se

Abstract

Medical texts can be difficult to understand for laymen, due to a frequent occurrence of specialised medical terms. Replacing these difficult terms with easier synonyms can, however, lead to improved readability. In this study, we have adapted a method for assessing difficulty of words to make it more suitable to medical Swedish. The difficulty of a word was assessed not only by measuring the frequency of the word in a general corpus, but also by measuring the frequency of substrings of words, thereby adapting the method to the compounding nature of Swedish. All words having a MeSH synonym that was assessed as easier, were replaced in a corpus of medical text. According to the readability measure LIX, the replacement resulted in a slightly more difficult text, while the readability increased according to the OVIX measure and to a preliminary reader study.

1 Introduction

Our health, and the health of our family and friends, is something that concerns us all. To be able to understand texts from the medical domain, e.g. our own health record or texts discussing scientific findings related to our own medical problems, is therefore highly relevant for all of us.

Specialised terms, often derived from latin or greek, as well as specialised abbreviations, are, however, often used in medical texts (Kokkinakis and Toporowska Gronostaj, 2006). This has the effect that medical texts can be difficult to comprehend (Keselman and Smith, 2012). Comprehending medical text might be particularly challenging for those laymen readers who are not used to looking up unknown terms while reading. A survey of

Swedish Internet users showed, for instance, that users with a long education consult medical information available on the Internet to a much larger extent than users with a shorter education (Findahl, 2010, pp. 28–35). This discrepancy between different user groups is one indication that methods for simplifying medical texts are needed, to make the medical information accessible to everyone.

Previous studies have shown that replacing difficult words with easier synonyms can reduce the level of difficulty in a text. The level of difficulty of a word was, in these studies, determined by measuring its frequency in a general corpus of the language; a measure based on the idea that frequent words are easier than less frequent, as they are more familiar to the reader. This synonym replacement method has been evaluated on medical English text (Leroy et al., 2012) as well as on Swedish non-medical text (Keskisärkkä and Jönsson, 2012). To the best of our knowledge, this method has, however, not previously been evaluated on medical text written in Swedish. In addition, as Swedish is a compounding language, laymen versions of specialised medical terms are often constructed by compounds of every-day Swedish words. Whether a word consists of easily understandable constituents, is a factor that also ought to be taken into account when assessing the difficulty of a word.

The aim of our study was, therefore, to investigate if synonym replacement based on term frequency could be successfully applied also on Swedish medical text, as well as if this method could be further developed by adapting it to the compounding nature of Swedish.

2 Background

The level of difficulty varies between different types of medical texts (Leroy et al., 2006), but studies have shown that even brochures intended

for patients, or websites about health issues, can be difficult to comprehend (Kokkinakis et al., 2012; Leroy et al., 2012). Bio-medical texts, such as medical journals, are characterised by sentences that have high informational and structural complexity, thus containing a lot of technical terms (Friedman et al., 2002). An abundance of medical terminology and a frequent use of abbreviations form, as previously mentioned, a strong barrier for comprehension when laymen read medical text. Health literacy is a much larger issue than only the frequent occurrence of specialised terms; an issue that includes many socio-economic factors. The core of the issue is, however, the readability of the text, and adapting word choice to the reader group (Zeng et al., 2005; Leroy et al., 2012) is a possible method to at least partly improve the readability of medical texts.

Semi-automatic adaption of word choice has been evaluated on English medical text (Leroy et al., 2012) and automatic adaption on Swedish non-medical text (Keskisärkkä and Jönsson, 2012). Both studies used synonym lexicons and replaced words that were difficult to understand with more easily understandable synonyms. The level of difficulty of a word was determined by measuring its frequency in a general corpus. The English study based its figures for word frequency on the number of occurrences of a word in Google's index of English language websites, while the Swedish study used the frequency of a word in the Swedish Parole corpus (Gellerstam et al., 2000), which is a corpus compiled from several sources, e.g. newspaper texts and fiction.

The English study used English WordNet as the synonym resource, and difficult text was transformed by a medical librarian, who chose easier replacements for difficult words among candidates that were presented by the text simplification system. Also hypernyms from semantic categories in WordNet, UMLS and Wiktionary were used, but as clarifications for difficult words (e.g. in the form: '*difficult word*, a kind of *semantic category*'). A frequency cut-off in the Google Web Corpus was used for distinguishing between easy and difficult words. The study was evaluated by letting readers 1) assess perceived difficulty in 12 sentences extracted from medical texts aimed at patients, and 2) answer multiple choice questions related to paragraphs of texts from the same resource, in order to measure actual difficulty. The

evaluations showed that perceived difficulty was significantly higher before the transformation, and that actual difficulty was significantly higher for one combination of medical topic and test setting.

The Swedish study used the freely available SynLex as the resource for synonyms, and one of the studied methods was synonym replacement based on word frequency. The synonym replacement was totally automatic and no cut-off was used for distinguishing between familiar and rare words. The replacement algorithm instead replaced all words which had a synonym with a higher frequency in the Parole corpus than the frequency of the original word. The effect of the frequency-based synonym replacement was automatically evaluated by applying the two Swedish readability measures LIX and OVIX on the original and on the modified text. Synonym replacement improved readability according to these two measures for all of the four studied Swedish text genres: newspaper texts, informative texts from the Swedish Social Insurance Agency, articles from a popular science magazine and academic texts.

For synonym replacement to be a meaningful method for text simplification, there must exist synonyms that are near enough not to change the content of what is written. Perfect synonyms are rare, as there is typically at least one aspect in which two separate words within a language differ; if it is not a small difference in meaning, it might be in the context in which they are typically used (Saeed, 1997). For describing medical concepts, there is, however, often one set of terms that are used by health professionals, whereas another set of laymen's terms are used by patients (Leroy and Chen, 2001; Kokkinakis and Toporowska Gronostaj, 2006). This means that synonym replacement could have a large potential for simplifying medical text, as there are many synonyms within this domain, for which the difference mainly lies in the context in which they are typically used.

The availability of comprehensive synonym resources is another condition for making it possible to implement synonym replacement for text simplification. For English, there is a consumer health vocabulary initiative connecting laymen's expressions to technical terminology (Keselman et al., 2008), as well as several medical termi-

Original	Med röntgen kan man se en ökad trabekulering, <i>osteoporos</i> samt pseudofrakturer.
Transformed	Med röntgen kan man se en ökad trabekulering, <i>benskörhet</i> samt pseudofrakturer.
Translated original	With X-ray, one can see an increased trabeculation, <i>osteoporosis</i> and pseudo-fractures.
Translated transformed	With X-ray, one can see an increased trabeculation, <i>bone-brittleness</i> and pseudo-fractures.

Table 1: An example of how the synonym replacement changes a word in a sentence.

nologies containing synonymic expressions, e.g. MeSH¹ and SNOMED CT². Swedish, with fewer speakers, also has fewer lexical resources than English, and although SNOMED CT was recently translated to Swedish, the Swedish version does not contain any synonyms. MeSH on the other hand, which is a controlled vocabulary for indexing biomedical literature, is available in Swedish (among several other languages), and contains synonyms and abbreviations for medical concepts (Karolinska Institutet, 2012).

Swedish is, as previously mentioned, a compounding language, with the potential to create words expressing most of all imaginable concepts. Laymen's terms for medical concepts are typically descriptive and often consist of compounds of words used in every-day language. The word *humerusfraktur* (*humerus fracture*), for instance, can also be expressed as *överarmsbenbrott*, for which a literal translation would be *upper-arm-bone-break*. That a compound word with many constituents occurring in standard language could be easier to understand than the technical terms of medical terminology, forms the basis for our adaption of word difficulty assessment to medical Swedish.

3 Method

We studied simplification of one medical text genre; medical journal text. The replacement method, as well as the main evaluation method, was based on the previous study by Keskisärkkä and Jönsson (2012). The method for assessing word difficulty was, however, further developed compared to this previous study.

As medical journal text, a subset of the journal *Läkartidningen*, the Journal of the Swedish Medical Association (Kokkinakis, 2012), was used.

¹www.nlm.nih.gov/mesh/

²www.ihtsdo.org

The subset consisted of 10 000 randomly selected sentences from issues published in 1996. As synonym lexicon, the Swedish version of MeSH was used. This resource contains 10 771 synonyms, near synonyms, multi-word phrases with a very similar meaning and abbreviation/expansion pairs (all denoted as *synonyms* here), belonging to 8 176 concepts.

Similar to the study by Keskisärkkä and Jönsson (2012), the Parole corpus was used for frequency statistics. For each word in the *Läkartidningen* subset, it was checked whether the word had a synonym in MeSH. If that was the case, and if the synonym was more frequently occurring in Parole than the original word, then the original word was replaced with the synonym. An example of a sentence changed by synonym replacement is shown in Table 1.

There are many medical words that only rarely occur in general Swedish, and therefore are not present as independent words in a corpus of standard Swedish, even if constituents of the words frequently occur in the corpus. The method used by Keskisärkkä and Jönsson was further developed to handle these cases. This development was built on the previously mentioned idea that a compound word with many constituents occurring in standard language is easier to understand than a rare word for which this is not the case. When neither the original word, nor the synonym, occurred in Parole, a search in Parole was therefore instead carried out for substrings of the words. The original word was replaced by the synonym, in cases when the synonym consisted of a larger number of substrings present in Parole than the original word. To insure that the substrings were relevant words, they had to consist of a least four characters.

Exemplified by a sentence containing the word *hemangiom* (*hemangioma*), the extended replacement algorithm would work as follows: The al-

gorithm first detects that *hemangiom* has the synonym *blodkärlstumör* (*blood-vessel-tumour*) in MeSH. It thereafter establishes that neither *hemangiom* nor *blodkärlstumör* is included in the Parole corpus, and therefore instead tries to find substrings of the two words in Parole. For *hemangiom*, no substrings are found, while four substrings are found for *blodkärlstumör* (Table 2), and therefore *hemangiom* is replaced by *blodkärlstumör*.

Word	1	2	3	4
hemangiom	-	-	-	-
blodkärlstumör	blod	kärl	blodkärl	tumör

Table 2: Example of found substrings

As the main evaluation of the effect of the synonym replacement, the two readability measures used by Keskiärrkkä and Jönsson were applied, on the original as well as on the modified text. LIX (läsbarhetsindex, readability measure) is the standard metric used for measuring readability of Swedish texts, while OVIX (ordvariationsindex, word variation index) measures lexical variance, thereby reflecting the size of vocabulary in the text (Falkenjack et al., 2013).

The two metrics are defined as follows (Mühlenbock and Johansson Kokkinakis, 2009):

$$\text{LIX} = \frac{O}{M} + \frac{L \cdot 100}{O}$$

Where:

- O = number of words in the text
- M = number of sentences in the text
- L = number of long words in the text (more than 6 characters)

$$\text{OVIX} = \frac{\log(O)}{\log(2 - \frac{\log(U)}{\log(O)})}$$

Where:

- O = number of words in the text
- U = number of unique words in the text

The interpretation of the LIX value is shown in Table 3, while OVIX scores ranging from 60 to 69 indicate easy-to-read texts (Mühlenbock and Johansson Kokkinakis, 2009).

LIX-value	Genre
less than 25	Children's books
25-30	Easy texts
30-40	Normal text/fiction
40-50	Informative texts
50-60	Specialist literature
more than 60	Research, dissertations

Table 3: The LIX-scale, from Mühlenbock and Johansson Kokkinakis (2009)

To obtain preliminary results from non-automatic methods, a very small manual evaluation of correctness and perceived readability was also carried out. A randomly selected subset of the sentences in which at least one term had been replaced were classified into three classes by a physician: 1) The original meaning was retained after the synonym replacement, 2) The original meaning was only slightly altered after the synonym replacement, and 3) The original meaning was altered more than slightly after the synonym replacement. Sentences classified into the first category by the physician were further categorised for perceived readability by two other evaluators; both with university degrees in non-life science disciplines. The original and the transformed sentence were presented in random order, and the evaluators were only informed that the simplification was built on word replacement. The following categories were used for the evaluation of perceived readability: 1) The two presented sentences are equally easy/difficult to understand, 2) One of the sentences is easier to understand than the other. In the second case, the evaluator indicated which sentence was easier.

4 Results

In the used corpus subset, which contained 150 384 tokens (26 251 unique), 4 909 MeSH terms for which there exist a MeSH synonym were found. Among these found terms, 1 154 were replaced with their synonym. The 15 most frequently replaced terms are shown in Table 4, many of them being words typical for a professional language that have been replaced with compounds of every-day Swedish words, or abbreviations that have been replaced by an expanded form.

The total number of words increased from 150 384 to 150 717 after the synonym replace-

Original term	(English)	Replaced with	(Literal translation)	n
aorta	(aorta)	kroppspulsåder	(body-artery)	34
kolestas	(cholestasis)	gallstas	(biliary-stasis)	33
angioödem	(angioedema)	angioneurotiskt ödem	(angio-neurotic-oedema)	29
stroke	(stroke)	slaganfall	(strike-seizure)	29
TPN	(TPN)	parenteral näring, total	(parenteral nutrition, total)	26
GCS	(GCS)	Glasgow Coma Scale	(Glasgow Coma Scale)	20
mortalitet	(mortality)	dödlighet	(deathrate)	20
ödem	(oedema)	svullnad	(swelling)	20
legitimation	(licensure)	licens	(certificate)	18
RLS	(RLS)	rastlösa ben-syndrom	(restless legs-syndrome)	18
anemi	(anemia)	blodbrist	(blood-shortage)	17
anhöriga	(family)	familj	(family)	17
ekokardiografi	(echocardiography)	hjärtultraljuds-undersökning	(heart-ultrasound-examination)	17
artrit	(arthritis)	ledinflammation	(joint-inflammation)	16
MHC	(MHC)	histokompatibilitetskomplex	(histocompatibility-complex)	15

Table 4: The 15 most frequently replaced terms. As the most frequent synonym (or synonym with most known substrings) is always chosen for replacement, the same choice among a number of synonyms, or a number of abbreviation expansions, will always be made. The column **n** contains the number of times the original term was replaced with this synonym.

ment. Also the number of long words (more than six characters) increased from 51 530 to 51 851. This resulted in an increased LIX value, as can be seen in Table 5. Both before and after the transformation, the LIX-value lies on the border between the difficulty levels of informative texts and non-fictional texts. The replacement also had the effect that the number of unique words decreased with 138 words, which resulted in a lower OVIX, also to be seen in Table 5.

For the manual evaluation, 195 sentences, in which at least one term had been replaced, were randomly selected. For 17% of these sentences, the original meaning was slightly altered, and for 10%, the original meaning was more than slightly altered. The rest of the sentences, which retained their original meaning, were used for measuring perceived readability, resulting in the figures shown in Table 6. Many replaced terms occurred more than once among the evaluated sentences. Therefore, perceived difficulty was also measured for a subset of the evaluation data, in which it was ensured that each replaced term occurred exactly once, by only including the sentence in which it first appeared. These subset figures (denoted *Unique* in Table 6) did, however, only differ marginally from the figures for the en-

tire set. Although there was a large difference between the two evaluators in how they assessed the effect of the synonym replacement, they both classified a substantially larger proportion of the sentences as easier to understand after the synonym replacement.

	LIX	OVIX
Original text	50	87.2
After synonym replacement	51	86.9

Table 5: LIX and OVIX before and after synonym replacement

5 Discussion

According to the LIX measure, the medical text became slightly more difficult to read after the transformation, which is the opposite result to that achieved in the study by Keskisärkkä and Jönsson (2012). Similar to this previous study, however, the text became slightly easier to read according to the OVIX measure, as the number of unique words decreased. As words longer than six characters result in a higher LIX value, a very plausible explanation for the increased LIX-value, is that short words derived from Greek or Latin have been replaced with longer compounds

Perceived effect of replacement	Evaluator 1 All (Unique)	Evaluator 2 All (Unique)
No difference	51% (52%)	29% (28%)
Easier	42% (42%)	54% (52%)
More difficult	7% (7%)	17% (21%)

Table 6: Results for the manual classification of perceived difficulty. Evaluator 1 classified 143 sentences and Evaluator 2 classified 140 sentences. The (Unique) column contains results when only the first occurrence of a replacement of a particular term is included. A binomial sign test (Newbold et al., 2003, p. 532) was performed on the effect of the replacement, with the null hypothesis that the probability of creating a more difficult sentence was equal to that of creating an easier one. This hypothesis could be rejected for both evaluators; when including all sentences and also when only including the (Unique) subset, showing that the differences were statistically significant ($p \ll 0.01$).

of every-day words. Replacing an abbreviation or an acronym with its expanded long form has the same effect. Expanding acronyms also increases the number of words per sentence, which also results in a higher LIX value.

Studies on English medical text indicate, however, that simple surface measures do not accurately reflect the readability (Zeng-Treitler et al., 2007; Wu et al., 2013), and user studies have been performed to construct readability measures better adapted to the domain of medical texts (Kim et al., 2007; Leroy and Endicott, 2012). Therefore, although the manual evaluation was very limited in scope, the results from this evaluation might give a better indication of the effects of the system. This evaluation showed that the perceived readability often improved with synonym replacement, although there were also replacements that resulted in a decrease of perceived readability. Further studies are required to determine whether these results are generalisable to a larger group of readers. Such studies should also include an evaluation of actual readability, using methods similar to those of Leroy et al. (2012). The cases, in which the synonym replacement resulted in a perceived decrease in readability should also be further studied. It might, for instance, be better to use a frequency cut-off for distinguishing between rare and frequent words, as applied by Leroy et al. (2012),

rather than always replacing a word with a more frequent synonym.

The manual evaluation also showed that the original semantic meaning had been at least slightly altered in almost a third of the sentences, which shows that the set of synonyms in Swedish MeSH might need to be adapted to make the synonyms suitable to use in a text simplification system. The replacements in Table 4 show three types of potential problems. First, there are also distant synonyms, as exemplified by *oedema* and *swelling*, where *oedema* means a specific type of swelling in the form of increased amount of liquid in the tissues, as opposed to e.g. increased amount of fat. Second, the MeSH terms are not always written in a form that is appropriate to use in running text, such as the term *parenteral nutrition, total*. Such terms need to be transformed to another format before they can be used for automatic synonym replacement. Third, although the abbreviations included in the manual evaluation were all expanded to the correct form, abbreviations within the medical domain are often overloaded with a number of different meanings (Liu et al., 2002). For instance, apart from being an acronym for *restless legs syndrome*, RLS can also mean *reaction level scale* (Cederblom, 2005). Therefore, in order to include abbreviations and acronyms in the synonym replacement method studied here, an abbreviation disambiguation needs to be carried out first (Gaudan et al., 2005; Savova et al., 2008). An alternative could be to automatically detect which abbreviations and acronyms that are defined in the text when they first are mentioned (Dannélls, 2006), and restrict the replacement method to those.

The sentence in Table 1 shows an example of a successful synonym replacement, replacing a word typically used by health professionals (*osteoporosis*) with a word typically used in everyday language (*bone-brittleness*). This sentence also gives an example of when not enough is replaced in the sentence for it to be easy to understand. Neither *trabeculation*, nor *pseudo-fractures*, are included in MeSH, which shows the importance of having access to comprehensive terminological resources for the method of synonym replacement to be successful. Extracting terms that are frequently occurring within the text genre that is to be simplified, but which are neither included in the used terminology, nor in a corpus

of standard language such as Parole, could be a method for finding candidates for expanding the terminological resources. Semi-automatic methods could be applied for finding synonyms to these new candidate terms, as well as to existing terms within the terminology for which no synonyms are provided (Henriksson et al., 2013).

Table 1 also exemplifies a further issue not addressed here, namely the frequent occurrence of inflected words in Swedish text. No morphologic normalisation, e.g. lemmatisation, was performed of the text that was to be simplified or of the terms in MeSH (e.g. normalising *pseudo-fractures* to *pseudo-fracture*). Such a normalisation would have the potential of matching, and thereby replacing, a larger number of words, but it would also require that the replaced word is inflected to match the grammatical form of the original word.

An alternative to using frequency in the Parole corpus, or occurrence of substrings in a word in Parole, for determining when a synonym is to be replaced, is to use the frequency in a medical corpus. That corpus then has to be targeted towards laymen, as word frequency in texts targeted towards health professionals would favour word replacements with words typical to the professional language. Examples of such patient corpora could be health related web portals for patients (Kokkinakis, 2011). However, as also texts targeted towards patients have been shown to be difficult to understand, the method of searching for familiar words in substrings of medical terms might be relevant for assessing word difficulty also if easy medical corpora would be used.

6 Future work

A number of points for future work have already been mentioned, among which evaluating the method on a large set of target readers has the highest priority. Adapting the method to handle inflected words, studying how near synonyms and ambiguity of abbreviations affect the content of the transformed sentences, as well as studying methods for semi-automatic expansion of terminologies, are other topics that have already been mentioned.

It might also be the case that what synonym replacements are suitable are dependent on the context in which a word occurs. Methods for adapting assessment of word difficulty to context have been studied within the Semeval-2012 shared task on

English lexical simplification (Specia et al., 2012), although it was shown that infrequent words are generally perceived as more difficult, regardless of context.

In addition to these points, it should be noted that we in this study have focused on one type medical text, i.e. medical journal text. As mentioned in the introduction, there is, however, another medical text type on which applying text simplification would also be highly relevant, namely health record text (Kvist and Velupillai, 2013; Kandula et al., 2010). The electronic health record is nowadays made available to patients via e-services in a number of countries, and there is also an on-going project constructing such a service in Sweden. Apart from health record text also containing many words derived from greek and latin, there are additional challenges associated with this type of text. As health record text is written under time pressure, it is often written in a telegraphic style with incomplete sentences and many abbreviations (Friedman et al., 2002; Aantaa, 2012). As was exemplified among the top 15 most frequently replaced words, abbreviations is one of the large problems when using the synonym replacement method for text simplification, as they are often overloaded with a number of meanings.

Future work, therefore, also includes the evaluation of synonym replacement on health record text. It also includes the study of writing tools for encouraging health professionals to produce text that is easier to understand for the patient, or at least easier to transform into more patient-friendly texts with methods similar to the method studied here (Ahltorp et al., 2013).

7 Conclusion

A method used in previous studies for assessing difficulty of words in Swedish text was further developed. The difficulty of a word was assessed not only by measuring the frequency of the word in a general corpus, but also by measuring the frequency of substrings of words, thereby adapting the method to the compounding nature of Swedish. The replacement was mainly evaluated by the two readability measures LIX and OVIX, showing a slightly decreased OVIX but a slightly increased LIX. A preliminary study on readers showed, however, an increased perceived readability after the synonym replacement. Studies on a larger reader group are required to draw any con-

clusions on the general effect of the method for assessment of word difficult. The preliminary results are, however, encouraging, showing that a method that replaces specialised words derived from latin and greek by compounds of every-day Swedish words can result in a increase of the perceived readability.

Acknowledgements

We would like to thank the three reviewers for many useful comments. This work was partly supported by a grant from the Vårdal Foundation

References

- Kirsi Aantaa. 2012. Mot patientvänligare epikriser, en kontrastiv undersökning [towards more patient friendly discharge letters, a contrastive study]. Master's thesis, Department of Nordic Languages, University of Turku.
- Magnus Ahltop, Maria Skeppstedt, Hercules Dalianis, and Maria Kvist. 2013. Using text prediction for facilitating input and improving readability of clinical text. *Stud Health Technol Inform*, 192:1149.
- Staffan Cederblom. 2005. *Medicinska förkortningar och akronymer (In Swedish)*. Studentlitteratur, Lund.
- Dana Dannélls. 2006. Automatic acronym recognition. In *Proceedings of the 11th conference on European chapter of the Association for Computational Linguistics (EACL)*.
- Johan Falkenjack, Katarina Heimann Mühlenbock, and Arne Jönsson. 2013. Features indicating readability in Swedish text. In *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013)*, pages 27–40.
- Olle Findahl. 2010. *Svenskarna och Internet*. .se.
- Carol Friedman, Pauline Kra, and Andrey Rzhetsky. 2002. Two biomedical sublanguages: a description based on the theories of zellig harris. *J Biomed Inform*, 35(4):222–35, Aug.
- S. Gaudan, H. Kirsch, and D. Rebholz-Schuhmann. 2005. Resolving abbreviations to their senses in medline. *Bioinformatics*, 21(18):3658–3664, September.
- M Gellerstam, Y Cederholm, and T Rasmark. 2000. The bank of Swedish. In *LREC 2000. The 2nd International Conference on Language Resources and Evaluation*, pages 329–333, Athens, Greece.
- Aron Henriksson, Mike Conway, Martin Duneld, and Wendy W. Chapman. 2013. Identifying synonymy between SNOMED clinical terms of varying length using distributional analysis of electronic health records. In *Proceedings of the Annual Symposium of the American Medical Informatics Association (AMIA 2013)*, Washington DC, USA.
- Sasikiran Kandula, Dorothy Curtis, and Qing Zeng-Treitler. 2010. A semantic and syntactic text simplification tool for health content. *AMIA Annu Symp Proc*, 2010:366–70.
- Karolinska Institutet. 2012. Hur man använder den svenska MeSHen (In Swedish, translated as: How to use the Swedish MeSH). http://mesh.kib.ki.se/swemesh/manual_se.html. Accessed 2012-03-10.
- Alla Keselman and Catherine Arnott Smith. 2012. A classification of errors in lay comprehension of medical documents. *Journal of Biomedical Informatics*, 45(6):1151–1163.
- Alla Keselman, Robert Logan, Catherine Arnott Smith, Gondy Leroy, and Qing Zeng-Treitler. 2008. Developing informatics tools and strategies for consumer-centered health communication. In *J Am Med Inform Assoc*, volume 15:4, pages 473–483.
- Robin Keskisarä and Arne Jönsson. 2012. Automatic text simplification via synonym replacement. In *Proceedings of Swedish Language Technology Conference 2012*.
- Hyeoneui Kim, Sergey Goryachev, Graciela Rosemblat, Allen Browne, Alla Keselman, and Qing Zeng-Treitler. 2007. Beyond surface characteristics: a new health text-specific readability measurement. *AMIA Annu Symp Proc*, pages 418–422.
- Dimitrios Kokkinakis and Maria Toporowska Gronostaj. 2006. Lay language versus professional language within the cardiovascular subdomain - a contrastive study. In *Proceedings of the 2006 WSEAS Int. Conf. on Cellular & Molecular Biology, Biophysics & Bioengineering*.
- Dimitrios Kokkinakis, Markus Forsberg, Sofie Johansson Kokkinakis, Frida Smith, and Joakim Öhlén. 2012. Literacy demands and information to cancer patients. In *Proceedings of the 15th International Conference on Text, Speech and Dialogue*, pages 64–71.
- Dimitrios Kokkinakis. 2011. Evaluating the coverage of three controlled health vocabularies with focus on findings, signs and symptoms. In *NEALT Proceedings Series*, editor, *NODALIDA*, volume 12, pages 27–31.
- Dimitrios Kokkinakis. 2012. The journal of the Swedish medical association - a corpus resource for biomedical text mining in Swedish. In *The Third Workshop on Building and Evaluating Resources for Biomedical Text Mining (BioTextM), an LREC Workshop*. Turkey.

- Maria Kvist and Sumithra Velupillai. 2013. Professional language in swedish radiology reports – characterization for patient-adapted text simplification. In *Scandinavian Conference on Health Informatics*, Copenhagen, Denmark, August.
- Gondy Leroy and Hsinchun Chen. 2001. Meeting medical terminology needs-the ontology-enhanced medical concept mapper. *IEEE Transactions on Information Technology in Biomedicine*, 5(4):261–270.
- Gondy Leroy and James E. Endicott. 2012. Combining nlp with evidence-based methods to find text metrics related to perceived and actual text difficulty. In *IHI*, pages 749–754.
- Gondy Leroy, Evren Eryilmaz, and Benjamin T. Laroya. 2006. Health information text characteristics. In *AMIA Annu Symp Proc*, pages 479–483.
- Gondy Leroy, James E Endicott, Obay Mouradi, David Kauchak, and Melissa L Just. 2012. Improving perceived and actual text difficulty for health information consumers using semi-automated methods. *AMIA Annu Symp Proc*, 2012:522–31.
- Hongfang Liu, Alan R Aronson, and Carol Friedman. 2002. A study of abbreviations in medline abstracts. *Proc AMIA Symp*, pages 464–8.
- Katarina Mühlenbock and Sofie Johansson Kokkinakis. 2009. Lix 68 revisited - an extended readability measure. In *Proceedings of Corpus Linguistics 2009*.
- Paul Newbold, William L. Carlson, and Betty Thorne. 2003. *Statistics for business and economics*. Prentice-Hall, Upper Saddle River, N. J., 5. ed. edition.
- John I. Saeed. 1997. *Semantics*. Blackwell Publishers, Oxford.
- Guergana K. Savova, Anni Coden, Igor L. Sominsky, Rie Johnson, Philip V. Ogren, Piet C. de Groen, and Christopher G. Chute. 2008. Word sense disambiguation across two domains: Biomedical literature and clinical notes. *Journal of Biomedical Informatics*, 41(6):1088–1100.
- Lucia Specia, Sujay Kumar Jauhar, and Rada Mihalcea. 2012. Semeval-2012 task 1: English lexical simplification. In **SEM, First Joint Conference on Lexical and Computational Semantics*, pages 347–355, Montréal, Canada.
- Danny T Y Wu, David A Hanauer, Qiaozhu Mei, Patricia M Clark, Lawrence C An, Jianbo Lei, Joshua Proulx, Qing Zeng-Treitler, and Kai Zheng. 2013. Applying multiple methods to assess the readability of a large corpus of medical documents. *Stud Health Technol Inform*, 192:647–51.
- Qing T. Zeng, Tony Tse, Jon Crowell, Guy Divita, Laura Roth, and Allen C. Browne. 2005. Identifying consumer-friendly display (cfd) names for health concepts. In *Proceedings of AMIA Annual Symposium*, pages 859–863.
- Qing Zeng-Treitler, Hyeoneui Kim, Sergey Goryachev, Alla Keselman, Laura Slaughter, and Catherine. A. Smith. 2007. Text characteristics of clinical reports and their implications for the readability of personal health records. *Medinfo*, 12(Pt 2):1117–1121.