

Automatic Summarization As Means Of Simplifying Texts, An Evaluation For Swedish

Christian Smith and Arne Jönsson

Santa Anna IT Research Institute AB

Linköping, Sweden

chrsm@ida.liu.se, arnjo@ida.liu.se

Abstract

We have developed an extraction based summarizer based on a word space model and PageRank and compared the readability of the resulting summaries with the original text, using various measures for Swedish and texts from different genres. The measures include among others readability index (LIX), nominal ratio (NR) and word variation index (OVIX). The measures correspond to the vocabulary load, idea density, human interest and sentence structure of the text and can be used to indicate the difficulty a reader might have in processing the text. The results show that the summarized texts are more readable, indicating that summarization can be used to reduce the effort to read a text.

1 Introduction

Many persons have, for various reasons, problems assimilating long complex texts. Not only persons with visual impairments or dyslexia, but also, for instance, those having a different mother tongue or persons in need of a quick summary of a text. In the project EasyReader we are developing an interactive tool for automatic summarization of texts from different genres.

Automatic summarization can be done in various ways. A common distinction is extract versus abstract summaries. An extract summary is created by extracting the most important sentences from the original text. An abstract summary on the other hand is a summary where the text has been broken down and rebuilt as a complete rewrite to convey a general idea of the original text. Furthermore, the summaries can be indicative (only providing keywords as central topics) or informative (content focused) (Firmin and Chrzanowski,

1999). The former might be more usable when a reader needs to decide whether or not the text is interesting to read and the latter when a reader more easily needs to get a grasp of the meaning of a text that is supposed to be read.

For various user groups it can also be beneficial if the text is easy to read and not only shorter. There are numerous measures of readability and readability for a number of different summarization techniques has been investigated by, for instance, Margarido et al. (2008). The difficulty of the text or the readability can also be measured by several automatic measures. Vadlapudi and Karagadda (2010) present an investigation on automatic evaluation of various aspects of readability for summaries. Readability has also been used, to, for instance, re-rank webpages to better suit a particular user (Newbold et al., 2010).

In this paper we will examine the readability of automatically created summaries by comparing them to the readability of the original full-length text. First, we describe the techniques used by the summarizer. Second, an overview of readability and some valid measures is presented. Third, we present results from the automatic readability evaluation of the generated summaries.

2 The summarizer

The summarizer used in our investigations is called COGSUM (Jönsson et al., 2008b; Jönsson et al., 2008a). It is an extraction based summarizer, using the word space model random indexing (RI), c.f. Hassel (2007) and a modified version of PageRank (Brin and Page, 1998).

2.1 The word space model

The word space model, or vector space model (Eldén, 2007), is a spatial representation of a word's meaning that can reduce the linguistic variability and capture semantically related concepts by taking into account the

positioning of words in a multidimensional space, instead of looking at only shallow linguistic properties. This facilitates the creation of summaries, since the positioning in the word space can be used to evaluate the different passages (words or sentences for instance) in relation to a document with regards to informational and semantic content.

Every word in a given context occupies a specific point in the space and has a vector associated to it that can be used to define its meaning.

Word spaces are constructed according to the distributional hypothesis and the proximity hypothesis. In the distributional hypothesis, words that occur in similar contexts have similar meanings so that a *word* is the sum of its contexts and the *context* is the sum of its words, where the *context* can be defined as the surrounding words or the entire document. The proximity hypothesis states that words close to each other in the word space have similar meaning while those far from each other have dissimilar meaning.

The word space is constructed from a matrix where text units are columns and the words in all text units are rows in the matrix. A certain entry in the matrix is nonzero iff the word corresponding to the row exists in the text unit represented by the column. The resulting matrix is very large and sparse which makes for the usage of techniques for reducing dimensionality. Latent Semantic Analysis is one such technique that, however, can be computationally expensive unless used with alternative algorithms (Gorrell, 2006).

2.2 Random Indexing (RI)

Random Indexing (Sahlgren, 2005; Kanerva, 1988) is a dimension reduction technique based on sparse distributed representations that provides an efficient and scalable approximate solutions to distributional similarity problems. The basic idea of Random Indexing is to accumulate context vectors based on the occurrence of words in contexts. This technique can be used with any type of linguistic context, is inherently incremental, and does not require a separate dimension reduction phase as for instance Latent Semantic Analysis.

Random Indexing can be described as a two-step operation:

Step 1 A unique d -dimensional *index vector* is assigned and randomly generated to each context (e.g. each document or each word).

These index vectors are sparse and high-dimensional. They consist of a small number, ρ , of randomly distributed +1s and -1s, with the rest of the elements of the vectors set to 0.

Step 2 *Context vectors* are produced on-the-fly. As scanning the text, each time a word occurs in a context (e.g. in a document, or within a sliding context window, w), that context's d -dimensional index vector is added to the context vector for the word. Words are thus represented by d -dimensional context vectors that are effectively the sum of the index vectors of all the contexts in which the word appears.

In COGSUM the vectors for whole sentences and the similarity between these and the average document vector are of interest. The average document vector is calculated by dividing the total document vector, which consists of the sum of all unique words' context vectors, with the number of unique words in the document, Equation 1.

$$\vec{doc} = \frac{1}{N} \sum_{i=1}^N \vec{w}_i \quad (1)$$

where N denotes the number of unique words.

The sentence vectors are then calculated by subtraction of the average document vector from the context vectors of the words in the sentence which are summed together and divided by the number of words in the sentence, Equation 2.

$$\vec{sent}_j = \frac{1}{S} \sum_{i=1}^S (\vec{w}_i - \vec{doc}) \quad (2)$$

where S denotes the number of words in sentence j .

2.3 PageRank

COGSUM uses the Weighted PageRank algorithm in conjunction to its RI-space to rank the sentences (Chatterjee and Mohan, 2007). PageRank is a graph-based ranking algorithm which originally was used to rank home pages automatically and objectively in the Google search engine (Brin and Page, 1998). To use PageRank for summaries we create an undirected fully connected graph where a vertex depicts a sentence in the current text and an edge between two different vertices is assigned a weight that depicts how similar these

are based on a cosine angle comparison of their meaning vectors, see Figure 1. As it is fully connected, all vertices are connected with each other.

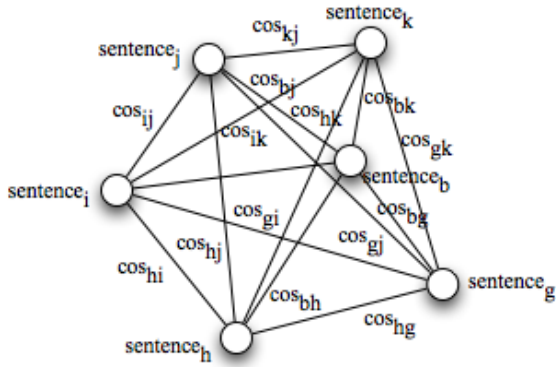


Figure 1: A simplified graph where sentences are linked and weighted according to the cosine values between them.

The algorithm rank ingoing and outgoing links to pages depending on the number of links as follows:

$$PR^W(s_i) = \frac{1-d}{N} + d \sum_{s_j \in In(s_i)} w_{ji} \frac{PR^W(s_j)}{\sum_{s_k \in Out(s_k)} w_{kj}} \quad (3)$$

where s_i is the sentence under consideration, $In(s_i)$ is the set of sentences that link to s_i , $Out(s_j)$ is the set of sentences that link from s_j and N is the total number of sentences. d is the damping factor.

The damping factor is originally set to account for the possibility of a surfer clicking a random web link when he gets bored (Brin and Page, 1998). With regards to the ranking of sentences, we see the damping factor as the possibility of a sentence containing some implicit information that a certain reader might consider more important at the time. The computation is carried out on all sentences iteratively until node weights converge.

Sentences with similar content will then contribute with positive support to each other. This does not exclusively depend on the number of sentences supporting a sentence, but also on the rank of the linking sentences. This means that a few high-ranked sentences provide bigger support than a greater number of low-ranked sentences. This leads to a ranking of the sentences by their importance to the document at hand and thus to a summary of desired length only including the most im-

portant sentences.

2.4 Achieving a summary

COGSUM, see Figure 2, takes as input only the text to be summarized along with a list of stop words (common function words such as prepositions). When the text has been processed using RI and PageRank, the most important sentences are extracted, for instance 30% of the original text, resulting in a condensed version of the original text with the most important information intact. Since all sentences are ranked, the length of the summary is easy to specify, in COGSUM this is implemented as a simple slider. COGSUM is designed for informative summaries, but it is also possible to have indicative summaries by clicking a "keywords" check box, see Figure 2.

It is important to note that the algorithm only takes the current document as total context and the information within the document, without any knowledge from an outside corpus (other than a list of stop words). This makes it highly portable to different domains, genres and languages (Mihalcea, 2004).

Evaluations of COGSUM with human users show that summaries produced by COGSUM are useful, considered informative enough and readable (Jönsson et al., 2008a). COGSUM has also been evaluated on gold standards for news texts and authority texts showing that it is better than another Swedish summarizer (SweSum, (Dalianis, 2000)) on authority texts and almost as good on news texts, texts that the other summarizer was especially adapted to handle (Gustavsson and Jönsson, 2010).

3 Readability

Generally, easy-to-read material is characterized by simple straightforward language without necessarily being simplistic or childish. Such material can be considered more readable and comprehensible by a person (Mühlenbock and Kokkinakis, 2009). Readability can, however, be seen from various angles. In part, it is the extent to which the reader can understand a written text, and the psychological processes involved within the reader. Here, focus lies on individual shortcomings with regards to perception and understanding of the written text and not on the text itself. Readability can also be seen as a measurable property of a given text. Then, the individual prerequisites

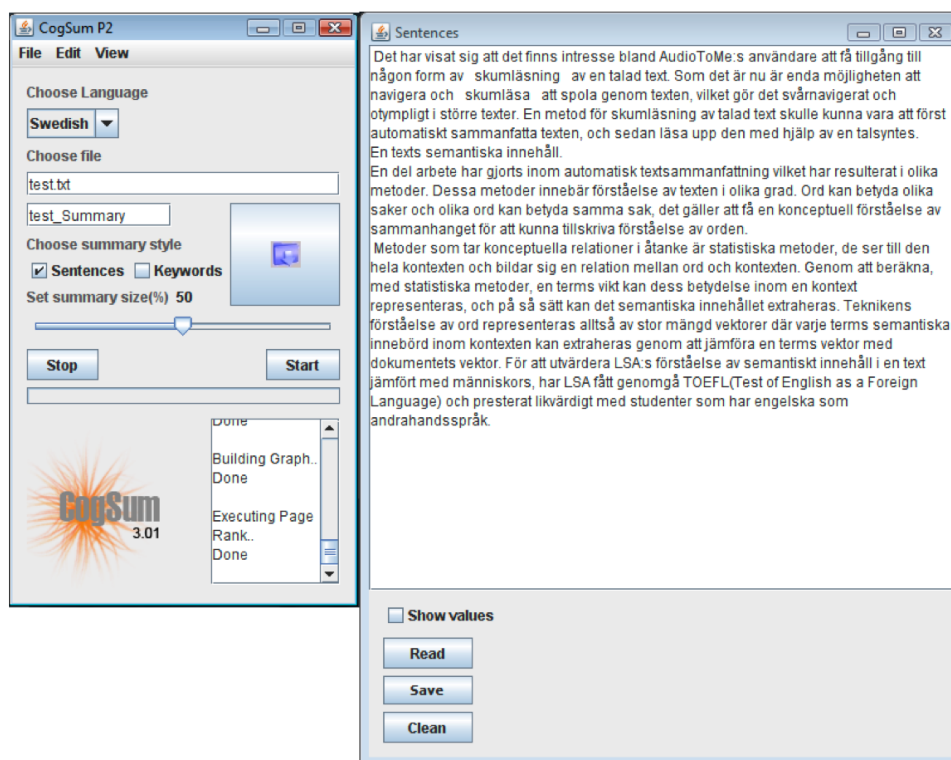


Figure 2: The COGSUM interface.

in terms of psychological abilities are often neglected.

Chall (1958) included the reader and concluded that there are four types of elements that seem to be significant for a readability criterion; vocabulary load, sentence structure, idea density and human interest. By mapping these psychological criteria of the individual with measurable readability properties of a text, several automatic measures of readability have been proposed.

3.1 Automatic measures

Automatic measures of readability for English is abundant, e.g. The Flesch Reading Ease Formula, Flesch-Kincaid, Dale-Chall, the Coleman-Liau test, Gunning Fog and SMOG (DuBay, 2004). For Swedish we have for instance: LIX, NR and OVIX (Mühlenbock and Kokkinakis, 2009; Rybing et al., 2010).

The Flesch Reading Easy score can be computed as:

$$Score = 206.835 - \left(1.015 \times \frac{n(w)}{n(s)}\right) - (84.6 \times ASW) \quad (4)$$

where $n(w)$ denotes the number of words, $n(s)$ the number of sentences and ASW the number of

syllables.

The measures correspond to how understandable a text is, e.g. a Flesch Reading Easy score between 70 and 80 is "Fairly Easy" which means that the text can easily be understood by a (U.S.) 7th grade student. The Flesch-Kincaid Grade level is a U.S. grade level version that normalises (4) to correspond to readability for students in various grades.

For Swedish, being an inflecting and compounding language, the readability index LIX (Björnsson, 1968) is almost exclusively used. LIX measures the number of words per sentence and also the number of long words (> 6 characters) in the text through the formula:

$$LIX = \frac{n(w)}{n(s)} + \left(\frac{n(words > 6 chars)}{n(w)} \times 100\right) \quad (5)$$

where $n(s)$ denotes the number of sentences and $n(w)$ the number of words.

Contrary to Flesch's original formula (and many of its modifications) the LIX formula does not consider syllables but instead word length. As LIX only considers ratios, sentence length and proportion of long words, it does not depend on text length.

A text with many long words and long sentences is considered more complex and therefore more difficult to read. A high LIX value indicates this, see Table 1.

Table 1: LIX-values for different genres, from Mühlenbock and Kokkinakis (2009)

LIX value	Text genre
-25	Children’s books
25-30	Easy texts
30-40	Normal text/fiction
40-50	Informative text
50-60	Specialist literature
> 60	Research, dissertations

In an effort to enhance the LIX formula, Mühlenbock and Kokkinakis (2009) have included an additional parameter called Extra Long Words (XLW). Extra Long Words are words with more than 14 characters and indicates a larger proportion of compounds of usually three or more stems, relatively common in Swedish. LIX has however been considered insufficient for a complete readability assessment of a text, since target groups of readers often are highly heterogeneous. By taking into account additional parameters, a better view of a text’s readability can be achieved. This is done by mapping Chall’s elements of readability to automatic measures (Mühlenbock and Kokkinakis, 2009). LIX is for instance, together with the amount of extra long words, mapped to *vocabulary load*.

Lexical variation or OVIX (word variation index) measures the ratio of unique tokens in a text and is used to indicate the *idea density*, in conjunction with the nominal ratio (NR). OVIX is calculated as:

$$OVIX = \frac{\log(n(w))}{\log(2 - \frac{\log(n(uw))}{\log(n(w))})} \quad (6)$$

where $n(w)$ denotes the number of words and $n(uw)$ the number of unique words. OVIX does not depend on text length (Lundberg and Reichenberg, 2009).

NR is calculated by dividing the number of nouns, prepositions and participles with the number of pronouns, adverbs and verbs:

$$NR = \frac{n(noun) + n(preposition) + n(participle)}{n(pronoun) + n(adverb) + n(verb)} \quad (7)$$

where $n(noun)$ denotes the number of nouns, $n(preposition)$ the number of prepositions, $n(participle)$ the number of participles, $n(pronoun)$ the number of pronouns, $n(adverb)$ the number of adverbs, and $n(verb)$ the number of verbs.

A higher NR indicates a more professional and stylistically developed text, while a lower value indicate more simple and informal language. In some contexts a low NR can indicate a narrative style, such as in children’s books. NR should not depend on text length.

The degree of *human interest* is measured simply through the proportion of proper nouns (PN) and by measuring the length of sentences (ASL), *sentence structure* can broadly be gathered (Mühlenbock and Kokkinakis, 2009).

4 Evaluation

We have evaluated summarized texts from a readability perspective by creating summaries of texts from different genres and compared their readability to the original text.

We use three types of texts representing three different genres:

- **DN.** Newspaper texts from the Swedish newspaper ”Dagens Nyheter”; ca 25,000 words divided in 130 articles.
- **FOF.** Popular science texts from the Swedish Magazine ”Forskning och Framsteg”; ca 20,000 words divided in 31 articles.
- **FOKASS.** Authority texts from the Swedish Social Insurance Administration (Sw. Försäkringskassan); ca 25,000 words from 2 brochures. The brochures were divided so that each chapter was an article resulting in a total of 35 ”articles”

The texts were summarized using COGSUM with a random index dimensionality, d , of 100, a focus window size, w , of 4 (2 left, 2 right) and $\rho = 4$, i.e. 2 positive 1:s and 2 negative 1:s, in line with Chatterjee and Mohan (2007). The texts were also stemmed using the snowball algorithm (Swedish) and stop words were removed. The PageRank damping factor was set to .85 (Brin and Page, 1998) and the number of iterations when the weights converged was below 50.

The texts were extracted from the concordances at Språkbanken (2011), except for the authority texts which were taken from the

Swedish Social Insurance Administration's web page (Försäkringskassan, 2011). They were summarized to different lengths (30%, 50% and 70%) and compared with the originals (100%) with regards to the different readability measures.

The summaries were evaluated using 7 measures, see Table 2. The values of the measures of the summaries were also compared to the full length texts using a paired-samples T-test.

5 Results

Table 2 shows the mean values for the various readability measures used on the different texts. Roughly, low values on any readability measure means that the text is more readable. The table also includes values on average word length (AWL).

The following significant differences were found:

DN For newspaper articles LIX got a lower score on all the summaries, ($p < .05$):

	Length	t(122)	Mean	SD
LIX	30%	-8.092	43.60	10.14
	50%	-9.147	45.21	8.70
	70%	-7.393	47.04	8.16
	100%		49.34	7.35

OVIX got a higher value for the 30% summary, ($p < .05$):

	Length	t(122)	Mean	SD
OVIX	30%	2.483	81.24	30.62
	100%		75.48	11.00

At 50% of the original text the words are also shorter on average (AWL) ($p < .05$):

	Length	t(122)	Mean	SD
AWL	50%	-3.4642	4.74	0.51
	100%		4.83	0.41

The sentences also became shorter (ASL) for all summarization lengths ($p < .05$):

	Length	t(122)	Mean	SD
ASL	30%	-4.817	16.12	5.09
	50%	-3.331	16.85	5.98
	70%	-5.115	17.04	4.27
	100%		18.00	3.91

FOF For popular science LIX was lower on all summarization lengths ($p < .05$):

	Length	t(30)	Mean	SD
LIX	30%	-6.933	53.65	9.84
	50%	-6.270	55.90	8.61
	70%	-5.327	57.57	8.36
	100%		59.92	7.81

At 50% and 70% OVIX got a lower score ($p < .05$):

	Length	t(30)	Mean	SD
OVIX	50%	-7.136	64.26	9.67
	70%	-6.017	66.26	8.89
	100%		69.24	7.94

At 30% and 50% we had lower average word length, ($p < .05$):

	Length	t(30)	Mean	SD
AWL	30%	-2.234	4.86	0.36
	50%	-2.465	4.89	0.27
	100%		4.94	0.23

We had a smaller proportion of extra long words for all summarization lengths ($p < .05$):

	Length	t(30)	Mean	SD
XLW	30%	-2.689	0.01	0.01
	50%	-2.458	0.01	0.01
	70%	-2.464	0.01	0.01
	100%		0.02	0.01

FOKASS Authority texts also displayed a lower LIX for all summarization lengths ($p < .05$):

	Length	t(34)	Mean	SD
LIX	30%	-8.497	46.28	12.76
	50%	-5.939	50.53	13.39
	70%	-4.642	52.92	13.09
	100%		55.46	13.00

OVIX was lower for 70% summarizations ($p < .05$):

	Length	t(34)	Mean	SD
OVIX	70%	-2.209	46.77	9.55
	100%		48.19	8.69

The sentences were longer at 50% and at 70% ($p < .05$):

	Length	t(34)	Mean	SD
ASL	50%	2.144	15.10	3.55
	70%	2.606	14.87	2.61
	100%		14.27	2.41

No significant differences could be observed in nominal ratio (NR) or proper nouns (PN) for any text genre or summarization length.

6 Discussion

A significantly lower LIX could be observed across the board of summaries, regardless of length and genre. This shows that the complexity of the text is reduced when the text is summarized by the summarizer.

A lower LIX presents together with the amount of extra long words the vocabulary load required to read the text (Mühlenbock and Kokkinakis, 2009). For popular science, this seems the most promi-

Length	TEXT	LIX		OVIX		NR		AWL		ASL		XLW		PN	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0,30	DN	43,60	10,14	<i>81,24</i>	30,62	1,39	0,72	4,76	0,62	16,12	5,09	0,02	0,02	0,09	0,07
	FOF	53,65	9,84	<i>65,29</i>	16,46	1,56	0,33	4,86	0,36	<i>17,46</i>	4,40	0,01	0,01	0,03	0,03
	FOKASS	46,28	12,76	<i>48,22</i>	14,17	1,23	0,62	4,84	0,63	<i>15,22</i>	4,30	0,04	0,03	0,02	0,03
0,50	DN	45,21	8,70	<i>74,41</i>	16,45	1,37	0,61	4,74	0,51	16,85	5,98	0,02	0,02	0,09	0,07
	FOF	55,90	8,61	64,26	9,67	1,59	0,36	4,89	0,27	<i>17,15</i>	3,32	0,01	0,01	0,03	0,02
	FOKASS	50,53	13,39	<i>47,96</i>	14,61	1,29	0,71	4,89	0,58	<i>15,10</i>	3,55	0,04	0,03	0,02	0,03
0,70	DN	47,04	8,16	<i>74,33</i>	12,74	1,37	0,54	4,79	0,44	17,04	4,27	0,02	0,02	0,09	0,06
	FOF	57,57	8,36	66,26	8,89	1,56	0,28	4,90	0,25	<i>17,01</i>	3,27	0,01	0,01	0,03	0,02
	FOKASS	52,92	13,09	46,77	9,55	1,25	0,54	4,90	0,51	<i>14,87</i>	2,61	0,04	0,03	0,01	0,02
1,00	DN	49,34	7,35	<i>75,48</i>	11,00	1,35	0,44	4,83	0,41	18,00	3,91	0,02	0,01	0,09	0,06
	FOF	59,92	7,81	<i>69,24</i>	7,94	1,53	0,24	4,94	0,23	<i>16,74</i>	2,82	0,02	0,01	0,03	0,02
	FOKASS	55,46	13,00	<i>48,19</i>	8,69	1,25	0,63	4,89	0,47	<i>14,27</i>	2,41	0,04	0,05	0,01	0,02

Table 2: Means of the measured readability scores, LIX, NR, and OVIX. AWL is the average length of the words in the text, ASL is the average sentence length, XLW is the proportion of words longer than 14 characters, and PN is the proportion of proper nouns in the text. Means that are significantly better than the original texts of the same genre are in bold, whereas means that are worse are in italics.

ment, as not only LIX but also the amount of extra long words decreased for all summarization lengths. Thus, for popular science, the vocabulary load decreased when articles were summarized.

OVIX also seems to be most effectively reduced for popular science texts when summarized, indicating that idea density is also reduced.

The average sentence length can be seen as a way of analyzing the structure of the sentence, without adopting syntactic parsing (Mühlenbock and Kokkinakis, 2009). This seems to be most prominent in newspaper articles. Newspaper texts had a high idea density from the start (by a high variation in words, OVIX) and a low LIX. They benefitted from summarization by getting a lower value on average sentence length, or sentence structure, for all summarization lengths.

Authority texts did not benefit as much from summarization, for all summarization lengths, as the other genres. OVIX was lower for 70% summaries but sentences got, for instance, longer for long summaries (50% and 70%).

No significant differences were found for any text on the amount of proper nouns (PN) and NR. A low NR might indicate a stylistically simple text such as narrative children’s books while a higher NR is more common in advanced texts and since the summarizer at this point does nothing to rewrite sentences, a change in NR is not to be expected.

To conclude; there seems to be a difference between different genres in how a summary is affecting the readability of the texts. Popular science seems to benefit most by being summarized, followed by newspaper articles. The vocabulary load and complexity of sentence structure can be lowered in newspaper articles and popular science,

where also idea density in addition can be lowered.

7 Summary and future research

We have shown that automatic summarization have a positive impact on readability for texts from different genres and with different summarization lengths. This shows that summarization can be used as a promising means to make a text more easy to read and may work well as a first step in an effort to make texts available with reduced difficulties across several target domains for different types of texts.

The evaluations were done using a number of automatic measures of readability. A next step is to conduct experiments with humans to gain further insights on readability of the summaries. For instance, a general problem with extract summarizations is that the sentences that are extracted can refer to a sentence that hasn’t been extracted, resulting in a fragmentation of the text that may increase the workload required by the reader.

Rewriting the text automatically based on syntactical properties of a target easy-to-read corpus, such as presented by Rybing et al. (2010), will probably further increase readability, and will also be investigated in future studies. Future research also include investigations on the interaction effects, found in this study, between various readability measures.

Acknowledgement

This research was partly supported by a research grant from The Swedish Post and Telecom Agency (PTS). The authors are grateful to Henrik Danielsson for many fruitful discussion and valuable help with the analyses.

References

- C.H. Björnsson. 1968. *Läsbarhet*. Stockholm: Liber.
- Sergey Brin and Lawrence Page. 1998. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107–117.
- J.S. Chall. 1958. *Readability: An appraisal of research and application*. Columbus, OH: Ohio State University Press. Reprinted 1974. Epping, Essex, England: Bowker Publishing Company.
- Nilhadri Chatterjee and Shiwali Mohan. 2007. Extraction-based single-document summarization using random indexing. In *Proceedings of the 19th IEEE international Conference on Tools with Artificial intelligence – (ICTAI 2007)*, pages 448–455.
- Hercules Dalianis. 2000. Swesum – a text summarizer for swedish. Technical Report TRITA-NA-P0015, IPLab-174, NADA, KTH, Sweden.
- William H. DuBay. 2004. *Smart language: Readers, Readability, and the Grading of Text*. Costa Mesa: Impact Information.
- Lars Eldén. 2007. *Matrix Methods in Data Mining and Pattern Recognition*. Society for Industrial & Applied Mathematics (SIAM).
- Thérese Firmin and Michael J Chrzanowski, 1999. *An Evaluation of Automatic Text Summarization Systems*, volume 6073, pages 325–336. SPIE.
- Försäkringskassan. 2011. Försäkringskassans website, January. <http://www.forsakringskassan.se>.
- Genevieve Gorrell. 2006. *Generalized Hebbian Algorithm for Dimensionality Reduction in Natural Language Processing*. Ph.D. thesis, Linköping University.
- Pär Gustavsson and Arne Jönsson. 2010. Text summarization using random indexing and pagerank. In *Proceedings of the third Swedish Language Technology Conference (SLTC-2010)*, Linköping, Sweden.
- Martin Hassel. 2007. *Resource Lean and Portable Automatic Text Summarization*. Ph.D. thesis, ISRN-KTH/CSC/A-07/09-SE, KTH, Sweden.
- Arne Jönsson, Mimi Axelsson, Erica Bergenholm, Bertil Carlsson, Gro Dahlbom, Pär Gustavsson, Jonas Rybing, and Christian Smith. 2008a. Skim reading of audio information. In *Proceedings of the The second Swedish Language Technology Conference (SLTC-08)*, Stockholm, Sweden.
- Arne Jönsson, Bjarte Bugge, Mimi Axelsson, Erica Bergenholm, Bertil Carlsson, Gro Dahlbom, Robert Krevers, Karin Nilsson, Jonas Rybing, and Christian Smith. 2008b. Using language technology to improve interaction and provide skim reading abilities to audio information services. In *Proceedings of eChallenges e-2008*, Stockholm, Sweden.
- Pentti Kanerva. 1988. *Sparse distributed memory*. Cambridge MA: The MIT Press.
- Ingvar Lundberg and Monica Reichenberg. 2009. *Vad är lättläst?* Socialpedagogiska skolmyndigheten.
- Paulo R. A. Margarido, Thiago A. S. Pardo, Gabriel M. Antonio, Vinícius B. Fuentes, Rachel Aires, Sandra M. Aluísio, and Renata P. M. Fortes. 2008. Automatic summarization for text simplification: Evaluating text understanding by poor readers. In *Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web*.
- Rada Mihalcea. 2004. Graph-based ranking algorithms for sentence extraction, applied to text summarization. In *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions*, ACLdemo '04, Morristown, NJ, USA. Association for Computational Linguistics.
- Katarina Mühlenbock and Sofie Johansson Kokkinakis. 2009. Lix 68 revisited – an extended readability measure. In *Proceedings of Corpus Linguistics*.
- Neil Newbold, Harry McLaughlin, and Lee Gillam. 2010. Rank by readability: Document weighting for information retrieval. In Hamish Cunningham, Allan Hanbury, and Stefan M. Rieger, editors, *Advances in Multidisciplinary Retrieval, First Information Retrieval Facility Conference, IRFC 2010, Vienna, Austria, May 31, 2010. Proceedings*, volume 6107 of *Lecture Notes in Computer Science*, pages 20–30. Springer.
- Jonas Rybing, Christian Smith, and Annika Silvervarg. 2010. Towards a rule based system for automatic simplification of texts. In *Proceedings of the third Swedish Language Technology Conference (SLTC-2010)*, Linköping, Sweden.
- Magnus Sahlgren. 2005. An Introduction to Random Indexing. *Methods and Applications of Semantic Indexing Workshop at the 7th International Conference on Terminology and Knowledge Engineering, TKE 2005*.
- Språkbanken. 2011. Concordances of språkbanken, January. <http://spraakbanken.gu.se/konk/>.
- Ravikiran Vadlapudi and Rahul Katragadda. 2010. On automated evaluation of readability of summaries: Capturing grammaticality, focus, structure and coherence. In *Proceedings of the NAACL HLT 2010 Student Research Workshop*.