Readability of Bangla News Articles for Children

Zahurul Islam and Rashedur Rahman
AG Texttechnology
Institut für Informatik
Goethe-Universität Frankfurt
zahurul@em.uni-frankfurt.de, kamol.sustcse@gmail.com

Abstract

Many news papers publish articles for children. Journalists use their experience and intuition to write these. They might not aware of readability of articles they write. There is no evaluation tool or method available to determine how appropriate these articles are for the target readers. In this paper, we evaluate difficulty of Bangla news articles that are written for children.

1 Introduction

News is the communication of selected information on current events (Shirky, 2009). This communication is shared by various mediums such as *print*, *online* and *broadcasting*. A *newspaper* is a printed publication that contains news and other informative articles. There are many *newspapers* that are also published online. Due to the rapid growth of internet use, it is very common that more people read news online nowadays than before. Newspapers try to target certain audience through different topics and stories. Children are also in their target audience. This target group is their future reader.

Nowadays children also read news online. One third of children in developed countries such as *Netherlands*, *United Kingdoms* and *Belgium* browse internet for news (De Cock, 2012; De Cock and Hautekiet, 2012). Another study by Livingstone et al. (2010) showed that one fourth of the British children between age of *nine* and *nineteen* look for news on the internet. The ratio could be similar in other developed countries where most of the citizen have access over the internet.

The number of internet users also increasing in developing countries such as Bangladesh and India. According to the English Wikipedia[1], more than *thirty three* million people in Bangladesh use internet and many of them read news online. Also the *Alexa index*[2] shows that three Bangla news sites are in the list of ten most visited websites from Bangladesh.

All newspapers contain a variety of sections. These sections are based on different news topics. Some of the them are specific to children. The news for children will vary linguistically and cognitively than news for adults. This characteristic is similar to the websites dedicated for children. De Cock and Heutekiet (2012) observed difficulties for children to navigate these websites. Readability of the texts is one of the reasons. There is no specific guideline for writing texts for this target group. Journalists use their experience and intuition while writing. However, a text that is very easy to understand for an adult reader could be very difficult for a child. This difficulty motivate children readers to skip the newspaper in future.

The readability of a text relates to how easily human readers can process and understand a text. There are many text related factors that influence the readability of a text. These factors include very simple features such as type face, font size, text vocabulary as well as complex features like grammatical conciseness, clarity, underlying semantics and lack of ambiguity. Nielsen (2010) recommended font size of 14 for young children and 12 for adults.

[1] http://en.wikipedia.org/wiki/Internet_in_Bangladesh
[2] http://en.wikipedia.org/wiki/Alexa_Internet

28th Pacific Asia Conference on Language, Information and Computation pages 309–317

Readability classification, is a task of mapping text onto a scale of readability levels. We explore the task of automatically classifying documents based on their different readability levels. As an input, this function operates on various statistics relating to different text features.

In this paper, we train a readability classification model using a corpus compiled from textbooks and features inherited from our previous works Islam et al. (2012; 2014) and features from Sinha et al. (2012). Later we use the model to classify Bangla news articles for children from different well-known news sources from Bangladesh and West Bengal.

The paper is organized as follows: Section 2 discusses related work. Section 3 describes cognitive model of children in terms of readability followed by an introduction of the training corpus and news articles in Section 4. The features used for classification are described in Section 5, and our experiments and results in Section 6 are followed by a discussion in Section 7. Finally, we present our conclusions in Section 8.

2 Related Work

Most of the text readability research works use texts for adult readers. Only few numbers of related work available that only focus on texts for children. De Belder and Moens (2010) perform a study that transfers a complex text into a simpler text so that the target text become easier to understand for children. They have focused on two types of simplification: *lexical* and *syntactic*. Two traditional readability formulas: *Flesch-Kincaid* (Kincaid et al., 1975) and *Dale-Chall* (Dale and Chall, 1948; Dale and Chall, 1995) are used to measure reading difficulty. De Cock and Heutekiet (2012) performed a usability study to analyze websites for children. The study uses texts from different websites published in *English* and *Dutch*. The usability experiment shows that *previous knowledge* of children play an important role to read and understand texts. They have used *Flesch-Kincaid* (Kincaid et al., 1975) to determine the difficulty level of English texts and a variation of the same formula for Dutch texts.

Both of the related work mentioned above use traditional readability formulas to measure text difficulty. However these traditional formulas have significant drawbacks. These formulas assume that texts do not contain noise and the sentences are always well-formed. However this is not the case always. Traditional formulas require significant sample sizes of text, they become unreliable for a text that contains less than 300 words (Kidwell et al., 2011). Si and Callan (2001), Peterson and Ostendorf (2009) and Feng et al. (2009) show that these traditional formulas are not reliable. These formulas are easy to implement, but have a basic inability to model the semantic of vocabulary usage in a context. The most important limitation is that these measures are based only on surface characteristics of texts and ignore deeper properties. They ignore important factors such as comprehensibility, syntactical complexity, discourse coherence, syntactic ambiguity, rhetorical organizations and propositional density of texts. Longer sentences are not always syntactically complex and counting the number of syllables of a single word does not show word difficulty. That is why, the validity of these traditional formulas for text comprehensibility is often suspect. Two recent works on Bangla texts use two of these traditional formulas. Das and Roychudhury (2004; 2006) show that readability measures proposed by Kincaid et al. (1975) and Gunning (1952) work well for Bangla. However, the measures were tested only for seven documents, mostly novels.

Since there are not many linguistic tools available for Bangla, researchers are exploring language independent and surface features to measure difficulty of Bangla texts. Recently, in our previous works, we proposed a readability classifier for Bangla using *information-theoretic* features (Islam et al., 2012; Islam et al., 2014). We have achieved an *F-Score* of 86.46% by combining these features with some lexical features. Sinha et al. (2012) proposed two readability models that are similar to classical readability measures for English. They conducted a user experiment to identify important structural parameters of Bangla texts. These measures are based on the average *word length* (WL), the *number of poly-syllabic words* and the *number of consonant–conjuncts*. According to their experimental results, *consonant–conjuncts* plays an important role in texts in terms of readability.

From the beginning of research on text readability, researchers proposed different measures for

English (Dale and Chall, 1948; Dale and Chall, 1995; Gunning, 1952; Kincaid et al., 1975; Senter and Smith, 1967; McLaughlin, 1969). Many commercial readability tools use traditional measures. Fitzsimmons et al. (2010) stated that the SMOG (McLaughlin, 1969) readability measure should be preferred to assess the readability of texts on health care.

Due to recent achievements in linguistic data processing, different linguistic features are now in the focus of readability studies. Islam et al. (2012) summarizes related work regarding language model-based features (Collins-Thompson and Callan, 2004; Schwarm and Ostendorf, 2005; Aluisio et al., 2010; Kate et al., 2010; Eickhoff et al., 2011), POS-related features (Pitler and Nenkova, 2008; Feng et al., 2009; Aluisio et al., 2010; Feng et al., 2010), syntactic features (Pitler and Nenkova, 2008; Barzilay and Lapata, 2008; Heilman et al., 2007; Heilman et al., 2008; Islam and Mehler, 2013), and semantic features (Feng et al., 2009; Islam and Mehler, 2013). Recently, Hancke et al. (2012) found that morphological features influence the readability of German texts.

Due to unavailability of linguistic resources for Bangla, we did not explore any of the linguistically motivated features. We have inherited features from Islam et al. (2012; 2014) and Sinha et al. (2012), these features achieve reasonable classification accuracy.

Children's reading skills is influenced by their cognitive ability. The following section describes children's cognitive model and text readability.

3 Text Readability and Children

Children start building their cognitive skills from an early age. They use their cognitive skills to perform different tasks in different environments. Kali (2009) stated that children refine their motor skills and start to be involved in different social games when they are 5 to 6 years of age. From age of 6 to 8, children start to expand their vision beyond their immediate surroundings. Children from 8 to 12 years of age acquire the ability to present different entities of the world using concepts and abstract representations. Children become more interested in social interactions in their teenage years.

Children learn to recognize alphabets prior they developed motor skills. This lead to develop their reading skills. Reading skills require two processes: *word decoding* and *comprehension*. *Word decoding* is a process of identifying a pattern of alphabets. Children must have the knowledge about these and their patterns. For example: it is impossible to recognise any word from any language without knowledge of alphabets of that language. A pattern of alphabets carry a semantic in their cognitive knowledge.

Comprehension is a process of extracting meaning from a sequence of words. The sequence of words follow an order. It could be impossible for children to understand a sentence where the order of the words is random. Therefore, *word order* plays an important role in text comprehension. Reading is different than understanding a picture, it extracts meaning from words that are separated by white spaces. The *comprehension* process is also influenced by the memory system.

The cognitive system of humans contains three different memories: *sensory memory*, *working memory* and *long-term memory* (Rayner et al., 2012). The *sensory store* contains raw, un-analyzed information very briefly. The ongoing cognitive process takes place in *working memory* and the *long-term memory* is the permanent storehouse of knowledge about the world (Kali, 2009). Older children sometimes are better where they simply retrieve a word from their memory while reading. A younger children might have to *sound out* of a novel word spelling. However they are also able to retrieve some of the familiar words. Children derive meaning of a sentence by combining words to form *propositions* then combine them get the final meaning. Some children might struggle to recognize words which make them unable to establish links between words. Children without this problem able to recognize words and derive meaning from a whole sentence. Generally, older children are better reader due to their working memory capacity where they can store more of a sentence in their memory as they are able to identify propositions in the sentence (De Beni and Palladino, 2000). Older children are able to comprehend more than younger children because of recognizing ability and more working memory (Kali, 2009). They also know more about the world and skilled to use appropriate reading strategies.

In summary, children become skilled reader as their working memories develop over time, extract propositions and combine them to understand the meaning of a sentence.

4 Data

The goal of this study is to asses difficulty of news articles that are aimed for children. The reading ability of children is very different than adult readers. The preceding section describes cognitive developments of children in terms of readability. A children who is 10 years old will have different reading capability than a children who is 15 years of old. That is why, a corpus that is categorized by the ages of children would be an ideal resource as training corpus. Duarte and Weber (2011) proposed different categories of children based on their ages. The categorized list is relevant with our study. However, our categorized list is still different than their one. The corpus is categorized as following age ranges:

- early elementary: $7 - 9$ years old
- readers: $10 - 11$ years old
- old children: $12 - 13$ years old
- teenagers: $14 - 15$ years old
- old teenagers: $16 - 18$ years old
- adults: above 18 years old

In this paper, we train a model using support vector machine (SVM). This technique requires a training corpus. We compile the training corpus from textbooks that have been using for teaching in different school levels in Bangladesh. The following subsections describe the training corpus and children news articles.

4.1 Training Corpus

The training corpus targets top four age groups described above. Textbooks from grade *two* to grade *ten* are considered as sources for corpus compilation. Generally, in Bangladesh children start going to schools when they are 6 years of old and finish the grade *ten* when they are fifteen (Arends-Kuenning and Amin, 2004). In our previous studies, Islam et al. (2012; 2014), we compile the corpus from the same source. However, the latest version is more cleaned and contains more documents. It contains texts from 54 textbooks. Table 1 shows the statistics of average *document length* (DL), average *sentence length* (SL) and average word length (WL). Textbooks were written using ASCII encoding which required to be converted into Unicode. The classification distinguishes four readability classes: *very easy*, *easy*, *medium* and *difficult*. Documents of (school) grade *two*, *three* and *four* are included into the class *very easy*. Class *easy* covers texts of grade *five* and *six*. Texts of grade *seven* and *eight* were subsumed under the class *medium*. Finally, all texts of grade *nine* and *ten* are belong to the class *difficult*.

Classes	Docs	Avg. DL	Avg. SL	Avg. WL
Very easy	234	88.28	7.46	5.27
Easy	113	150.46	9.09	5.27
Medium	201	197.08	10.35	5.47
Difficult	113	251.30	12.19	5.66

Table 1: The Training Corpus.

4.2 News Articles

The goal of this paper is observing children news articles in Bangla on the basis of difficulty levels. As an Indo-Aryan language Banga is spoken in Southeast Asia, specifically in present day Bangladesh and the Indian states of West Bengal, Assam, Tripura and Andaman and on the Nicobar Islands. With nearly 250 million speakers (Karim et al., 2013), Bangla is spoken by a large speech community. However, due to lack of linguistic resources Bangla is considered as a low-resourced language.

We collected children news articles from four popular news sites from Bangladesh and one from West Bengal. The sites are: *Banglanews24*[3], *Bdnews24*[4], *Kaler kantho*[5], *Prothom alo*[6] and *Ichchhamoti*[7]. *Banglanews24*, *Bdnews24* and *Ichchhamoti* publish online only. In contrast, *Kalerkantho* and *Prothomalo* publish as printed newspapers and online. These newspapers publish weekly featured articles for children. We have collected 50 fea-

[3]www.banglanews24.com
[4]www.bangla.bdnews24.com
[5]www.kalerkantho.com
[6]www.prothomalo.com
[7]http://www.ichchhamoti.in/

tured articles from each of the sites and pre-process in similar way as the training corpus. However, the news articles are already written in Unicode and cover different topics ranges from *family*, *society*, *science* and *history* to *sports*. Table 2 shows different statistics of news articles.

News sites	Average DL	Average. SL	Average WL
Banglanews24	50.14	9.48	5.04
Bdnews24	62.66	9.82	4.91
Kaler kantho	53.08	8.90	4.89
Prothom alo	47.92	9.15	4.89
Ichchhamoti	105.50	11.86	4.66

Table 2: Statistics of news articles.

5 Feature Selection

A limited number of related works available that deal texts from Bangla. All of them are limited into traditional readability formulas, lexical and information-theoretic features. Any of features do not require any linguistic pre-processing. The following subsections describe feature selection in detail.

5.1 Lexical Features

We inherited a list of lexical features from our previous study Islam et al. (2014). Lexical features are very cheap to compute and shown useful for different text categorizing tasks. Average SL and average WL are two of most used features for readability classification. Recently, Learning (2001) showed that these are the two most reliable measures that affect readability of texts. The average SL is a quantitative measure of syntactic complexity. In most cases, the syntax of a longer sentence is difficult than the syntax of a shorter sentence. However, children of a lower grade level are not aware of syntax. A long word that contains many syllables is morphologically complex and leads to comprehension problems (Harly, 2008). Generally, most of the frequent words are shorter in length. These frequent words are more likely to be processed with a fair degree of automaticity. This automaticity increases reading speed and free-memory for higher level meaning building (Crossley et al., 2008).

Our previous study, Islam et al. (2014) also listed different *type token ratio* (TTR) formulas. The TTR indicates lexical density of texts, a higher value of it reflects the diversification of the vocabulary of a text. The diversification causes difficulties for children. In a diversified text, synonyms may be used to represent similar concepts. Children face difficulties to detect relationship between synonyms (Temnikova, 2012).

5.2 Information-Theoretic features

Nowadays, researchers exploring uncertainty based features from the field of *information theory* to measure complexity in natural languages (Febres et al., 2014). Information theory studies statistical laws of how information can be optimally coded (Cover and Thomas, 2006). The entropy rate plays an important role in human communication in general (Genzel and Charniak, 2002; Levy and Jaeger, 2007). The rate of information transmission per second in a human speech conversation is roughly constant, that is, transmitting a constant number of bits per second or maintaining a constant entropy rate. The entropy of a random variable is related to the difficulty of correctly guessing the value of the corresponding random variable. In our previous studies, Islam et al. (2012; 2014) and Islam and Mehler (2013) use different information-theoretic features for text readability classification. Our hypothesis was that the higher the entropy, the less readable the text along the feature represented by the corresponding random variable. We have inherited seven information-theoretic features from our previous studies.

5.3 Readability Models for Bangla

Recently, Sinha et al. (2012) proposed few computational models that are similar to the traditional English readability formulas. A user study was performed to evaluate their performance. We also inherited two of their best performing models:

$$Model3 = -5.23 + 1.43 * AWL + .01 * PSW \quad (1)$$

$$Model4 = 1.15 + .02 * JUK - .01 * PSW30 \quad (2)$$

In their models, they use structural parameters such as average WL, *number of jukta-akshars* (JUK) or consonant-conjuncts, *number of polysyllabic words* (PSW). The PSW30 shows that normalized value of PSW over 30 sentences.

Features	Accuracy	F-Score
Model 3	56.61%	49.13%
Model 4	56.38%	52.51%
Together	66.27%	65.67%

Table 3: Performance of Bangla readability models proposed by Sinha et al. (Sinha et al., 2012).

In this paper, we use 20 features to generate feature vectors for the classifier. The following section describes our experiments and results on training corpus and news articles.

6 Experiments and Results

In order to find the best performing training model, we use 20 features from Islam et al. (2012; 2014) and Sinha et al. (2012). Note that hundred data sets were randomly generated where 80% of the corpus was used for training and remaining 20% for evaluation. The weighted average of *Accuracy* and *F-score* is computed by considering results of all data sets. We use the SMO (Platt, 1998; Keerthi et al., 2001) classifier model implemented in WEKA (Hall et al., 2009) together with the Pearson VII function-based universal kernel PUK (Üstün et al., 2006).

6.1 Training Model

The traditional readability formulas that were proposed for English texts do not work for Bangla texts (Islam et al., 2012; Islam et al., 2014; Sinha et al., 2012). That is why, we did not explore any of the traditional formulas.

At first we build a classifier using two readability models from Sinha et al(2012). The output of these models are used as input for the readability classifier. Table 3 shows the evaluation results. The classification accuracy is little over than 66%. In our previous study Islam et al. (2014) found better classification accuracy using these features. However, the corpus is slightly different. The latest version of the corpus contains more documents for *easy* readability class. The classifier miss-classifies documents from this class mostly. The classifier labeled many of the documents from this readability class as *very easy*. Miss-classification of documents from other readability classes are also observed.

Table 4 shows the performance of features proposed in our previous study Islam et al. (2014).

Features	Accuracy	F-Score
Average SL	61.53%	55.21%
TTR (sentence)	47.32%	41.31%
TTR (document)	53.84%	52.61%
Average DW (sentence)	54.69%	55.28%
Number DW (document)	62.56%	60.12%
Avg. WL	44.63%	40.82%
Corrected TTR	59.38%	54.31%
Köhler TTR	54.61%	49.61%
Log TTR	47.49%	43.30%
Root TTR	60.76%	52.49%
Deviation TTR	52.32%	47.83%
Word prob.	60.76%	54.49%
Character prob.	50.00%	47.13%
WL prob.	51.58%	46.40%
WF prob.	52.30%	47.80%
CF prob.	60.76%	52.18%
SL and WL prob.	62.30%	59.74%
SL and DW prob.	66.92%	63.09%
18 features proposed by Islam et al. (2014)	85.60%	84.46%

Table 4: Performance of features proposed by Islam et al. (2014).

The classification accuracy also drops. The classifier also suffer to classify documents from *easy* readability class correctly. However, information-transmission based features (i.e., SL and WL prob. and SL and DW prob.) are the best performing features. Therefore, a text with higher average SL become more difficult when it contains more difficult words or more longer words.

The classification F-Score rises to 87.87 when we combine features from Islam et al. (2014) and Sinha et al. (Sinha et al., 2012).

6.2 News Articles Classification

Total 250 children news articles are collected as candidate news articles for classification. We consider the whole training corpus in order to build a training model. The training model is used to classify the candidate news articles. Among all articles, 160 articles are labeled as *very easy* and 18 articles as *easy*. Only 2 articles are labeled as *difficult* and remaining 60 articles are labeled as *medium*. Figure 1 shows classification results. More than 60% of news articles from newspapers are classified as *very easy*. However, the amount drops below 20% for the articles from *Icchamoti* children magazine. Also articles labeled as *difficult* belong to this magazine. The evaluation shows that, among all of the newspapers, news from *Banglanews24* are more suitable for children. Most of articles from that site belong to *very*

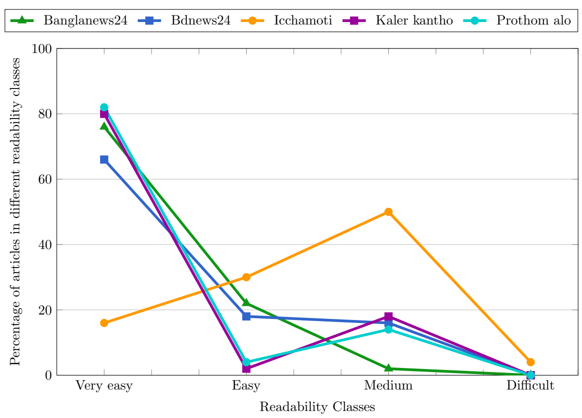

Figure 1: Classification of Bangla news articles for children.

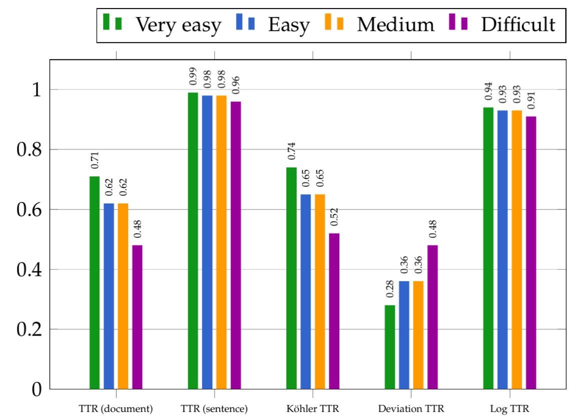

Figure 2: Observation of different TTR formulas in classified news articles.

easy and *easy* readability class.

Apart from the classification of children news articles we are also interested in behavior of different features in classified articles. The following section describes from interesting observation we notice.

7 Observation

Articles from Ichchhamoti has the lowest average WL. But, have higher values for average DW and average SL. Two of the articles from this site are labeled as *difficult*. This labeling could be influenced by average DW and average SL. Documents from training corpus have higher average WL.

Among the lexical features different TTRs have been considered to measure text difficulty (Islam et al., 2014). An article with a higher TTR value supposed to be difficult that an article with a lower TTR value (See Section 5.1). However, we observe different behavior of TTR formulas. Figure 2 shows the behaviour of different TTR formulas in classified articles. The average TTR value of articles from *very easy* readability class is higher than the average TTR value of articles from higher difficulty classes. Article length could be the reason of this irregularity. Articles from higher difficulty classes are longer and contain more words.

We also observed that some articles which have lower average SL, but labeled as *medium*. In contrast, some articles that have higher average SL, but labeled as *very easy* or *easy*. We randomly choose such articles and observe average SL. The average SL of articles belong to *medium* is 7.40 and the average SL of articles belongs to *easy* or *very easy* is 12.08. However, articles that are labeled as *medium* have higher average *word entropy* than articles that are labeled as *easy* or *very easy*. This shows that different type of features should be considered together to build a readability classifier.

8 Conclusion

In this paper, our goals was to examine the difficulty levels of news articles targeting children. Therefore we build a readability classifier that is able to classify the corresponding news articles into different difficulty levels. Children news articles are cognitively and linguistically different than articles for adult readers. A readability classifier trained on a textbooks corpus is able to classify these articles. Although linguistically motivated features could capture linguistic properties of news articles. Lexical features and features related to information density also have good predictive power to identify text difficulties. The classification results show that candidate articles are appropriate for children. This study also validate that features in our previous study Islam et al. (2014) and features proposed by Sinha et al. (Sinha et al., 2012) are useful for Bangla text readability analysis.

There are many languages in the world which lack a readability measurement tool. A readability classifier for these language could be built by using the features proposed in our previous study Islam et al.

(2014).

9 Acknowledgments

We like to thank Prof. Dr. Alexander Mehler for arranging money to travel the conference. We also like to thank the anonymous reviewers for their helpful comments. This work is funded by the LOEWE Digital-Humanities project at Goethe-University Frankfurt.

References

Ra Aluisio, Lucia Specia, Caroline Gasperin, and Carolina Scarton. 2010. Readability assessment for text simplification. In *NAACL-HLT 2010: The 5th Workshop on Innovative Use of NLP for Building Educational Applications*.

Mary Arends-Kuenning and Sajeda Amin. 2004. School incentive programs and childrens activities: The case of bangladesh. *Comparative Education Review*, 48(3):295–317.

Regina Barzilay and Mirella Lapata. 2008. Modeling local coherence: An entity-based approach. *Computational Linguistics*, 21(3):285–301.

Kevyn Collins-Thompson and James P Callan. 2004. A language modeling approach to predicting reading difficulty. In *HLT-NAACL*.

Thomas M. Cover and Joy A. Thomas. 2006. *Elements of Information Theory*. Wiley-Interscience, Hoboken.

Scott A Crossley, Jerry Greenfield, and Danielle S McNamara. 2008. Assessing text readability using cognitively based indices. *Tesol Quarterly*, 42(3):475–493.

Edgar Dale and Jeanne S. Chall. 1948. A formula for predicting readability. *Educational Research Bulletin*, 27(1):11–20+28.

Edgar Dale and Jeanne S. Chall. 1995. *Readability Revisited: The New Dale-Chall Readability formula*. Brookline Books.

Sreerupa Das and Rajkumar Roychoudhury. 2004. Testing level of readability in bangla novels of bankim chandra chattopodhay w.r.t the density of polysyllabic words. *Indian Journal of Linguistics*, 22:41–51.

Sreerupa Das and Rajkumar Roychoudhury. 2006. Readabilit modeling and comparison of one and two parametric fit: a case study in bangla. *Journal of Quantative Linguistics*, 13(1).

Jan De Belder and Marie-Francine Moens. 2010. Text simplification for children. In *Prroceedings of the SIGIR workshop on accessible search systems*, pages 19–26.

Rossana De Beni and Paola Palladino. 2000. Intrusion errors in working memory tasks: Are they related to reading comprehension ability? *Learning and Individual Differences*, 12(2):131–143.

Rozane De Cock and Eva Hautekiet. 2012. Childrens news online: Website analysis and usability study results (the united kingdom, belgium, and the netherlands). *Journalism and Mass Communication*, 2(12):1095–1105.

Rozane De Cock. 2012. Children and online news: a suboptimal relationship. quantitative and qualitative research in flanders. *E-youth: Balancing between opportunities a risks*.

Sergio Duarte Torres and Ingmar Weber. 2011. What and how children search on the web. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pages 393–402. ACM.

Carsten Eickhoff, Pavel Serdyukov, and Arjen P. de Vries. 2011. A combined topical/non-topical approach to identifying web sites for children. In *Proceedings of the fourth ACM international conference on Web search and data mining*.

Gerardo Febres, Klaus Jaffé, and Carlos Gershenson. 2014. Complexity measurement of natural and artificial languages. *Complexity*.

Lijun Feng, Noémie Elhadad, and Matt Huenerfauth. 2009. Cognitively motivated features for readability assessment. In *Proceedings of the 12th Conference of the European Chapter of the ACL*.

Lijun Feng, Martin Janche, Matt Huenerfauth, and Noémie Elhadad. 2010. A comparison of features for automatic readability assessment. In *The 23rd International Conference on Computational Linguistics (COLING)*.

PR Fitzsimmons, BD Michael, JL Hulley, and GO Scott. 2010. A readability assessment of online parkinsons disease information. *The Journal of the Royal College of Physicians of Edinburgh*, 40:292–296.

Dimitry Genzel and Eugene Charniak. 2002. Entropy rate constancy in text. In *Proceedings of the 40st Meeting of the Association for Computational Linguistics (ACL 2002)*.

Robert Gunning. 1952. *The Technique of clear writing*. McGraw-Hill; Fourh Printing Edition.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA data mining software: an update. *ACM SIGKDD Explorations*, 11(1):10–18.

Julia Hancke, Sowmya Vajjala, and Detmar Meurers. 2012. Readability classification for german using lexical, syntactic and morphological features. In *24th International Conference on Computational Linguistics (COLING), Mumbai, India*.

Trevor A. Harly. 2008. *The Psychology of Language*. Psychology Press, Taylor and Francis Group.

Michael Heilman, Kevyn Collins-Thompson, and Maxine Eskenazi. 2007. Combining lexical and grammatical features to improve readavility measures for first and second language text. In *Proceedings of the Human Language Technology Conference*.

Michael Heilman, Kevyn Collins-Thompson, and Maxine Eskenazi. 2008. An analysis of statistical models and features for reading difficulty prediction. In *Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications (EANL)*.

Zahurul Islam and Alexander Mehler. 2013. Automatic readability classification of crowd-sourced data based on linguistic and information-theoretic features. In *14th International Conference on Intelligent Text Processing and Computational Linguistics*.

Zahurul Islam, Alexander Mehler, and Rasherdur Rahman. 2012. Text readability classification of textbooks of a low-resource language. In *Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation.*

Zahurul Islam, Md Rashedur Rahman, and Alexander Mehler. 2014. Readability classification of bangla texts. In *Computational Linguistics and Intelligent Text Processing*, pages 507–518. Springer.

Robert V. Kali. 2009. *Children and Their Development*. Pearson Education.

MA Karim, M Kaykobad, and M Murshed. 2013. *Technical Challenges and Design Issues in Bangla Language Processing*. IGI Global.

Rohit J. Kate, Xiaoqiang Luo, Siddharth Patwardhan, Martin Franz, Radu Florian, Raymond J. Mooney, Salim Roukos, and Chris Welty. 2010. Learning to predict readability using diverse linguistic features. In *23rd International Conference on Computational Linguistics (COLING 2010)*.

S.S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy. 2001. Improvements to Platt's SMO algorithm for SVM classifier design. *Neural Computation*, 13(3):637–649.

Paul Kidwell, Guy Lebanon, and Kevyn Collins-Thompson. 2011. Statistical estimation of word acquisition with application to readability prediction. *Journal of the American Statistical Association*, 106(493):21–30.

J. Kincaid, R. Fishburne, R. Rodegers, and B. Chissom. 1975. Derivation of new readability formulas for Navy enlisted personnel. Technical report, US Navy, Branch Report 8-75, Cheif of Naval Training.

Renaissance Learning. 2001. The ATOS readability formula for books and how it compares to other formulas. *Madison, WI: School Renaissance Institute*.

Roger Levy and T. Florian Jaeger. 2007. Speakers optimize information density through syntactic reduction. *Advances in neural information processing systems*, pages 849–856.

Sonia Livingstone, Leslie Haddon, Anke Görzig, and Kjartan Ólafsson. 2010. Risks and safety for children on the internet: the uk report. *Politics*, 6(2010):1.

G. Harry McLaughlin. 1969. SMOG grading – a new readability formula. *Journal of Reading*, 12(8):639–646.

Jakob Nielsen. 2010. Children's websites: Usability issues in designing for kids. *Jakob Nielsens Alertbox*.

Sarah E. Petersen and Mari Ostendorf. 2009. A machine learning approach to reading level assesment. *Computer Speech and Language*, 23(1):89–106.

Emily Pitler and Ani Nenkova. 2008. Revisiting readability: A unified framework for predicting text quality. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

John C. Platt. 1998. *Fast training of support vector machines using sequential minimal optimization*. MIT Press.

Keith Rayner, Alexander Pollatsek, Jane Ashby, and Charles Clifton Jr. 2012. *Psychology of Reading*. Psychology Press.

Sarah E. Schwarm and Mari Ostendorf. 2005. Reading level assessment using support vector machines and statistical language models. In *the Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics(ACL 2005)*.

R.J. Senter and E. A. Smith. 1967. Automated readability index. Technical report, Wright-Patterson Air Force Base.

Clay Shirky. 2009. *Here comes everybody: How change happens when people come together*. Penguin UK.

Luo Si and Jamie Callan. 2001. A statistical model for scientific readability. In *Tenth International Conference on Information and Knowledge Management*.

Manjira Sinha, Sakshi Sharma, Tirthankar Dasgupta, and Anupam Basu. 2012. New readability measures for bangla and hindi texts. In *COLING (Posters)*, pages 1141–1150.

Irina Temnikova. 2012. *Text Complexity and Text Simplification in the Crisis Management Domain*. Ph.D. thesis, University of Wolverhampton.

B. Üstün, W.J. Melssen, and L.M.C. Buydens. 2006. Facilitating the application of support vector regression by using a universal Pearson VII function based kernel. *Chemometrics and Intelligent Laboratory Systems*, 81(1):29–40.