Readability Evaluation for Ukrainian Medicine Corpus (UKRMED)

Olga Cherednichenko and Olga Kanishcheva

National Technical University "Kharkiv Polytechnic Institute", 2, Kyrpychova str., Kharkiv, 61002, Ukraine

Abstract

In our work, we decided to demonstrate how to work different readability formulas on our Ukrainian-language corpus (UKRMED) of medical texts. UKRMED contains three types of texts in the medical domain divided by their complexity: "Complex texts", "Moderate texts", and "Simple texts". This research aims to (1) demonstrate the use of the most commonly used readability formulas on written health information in Ukrainian, (2) compare and contrast these different formulas to various texts (simple, complex, and moderate), (3) research different medical text features which will be used for text simplification and classification medical texts and (4) prepare recommendations for using these formulas to the evaluation of readability medical texts in Ukrainian.

Keywords 1

Text simplification, readability formulas, reading indexes, medicine text corpus, Ukrainian

1. Introduction

The perception of the text is very important when it comes to a special domain (for example, medicine, military science, mathematics, etc.) or a text in a foreign language. The task of assessing the text complexity is often set in a general meaning [1]. It is necessary to simplify the text in order to people whose education level is insufficient (for example, children) or who are not native speakers can perceive such text easier. However, in ordinary life, we often have to deal with texts that are hard to perceive even by educated native speakers. For example, texts on medical topics, such as official medical protocols, drug descriptions, medical records, etc. The situation is aggravated by a huge increase in information on the Internet. Internet users look for information on a medical topic and often read low-quality but clearly written texts. Blogs, forums, posts on social networks are becoming a source of information for many people. People do not read official medical literature because of the difficulty in perceiving such texts. Accordingly, we began our study of the complexity of medical texts in Ukrainian in [2].

A text corpus is an important resource for learning a language. In our research, we are faced with a lack of text resources in the Ukrainian language. This is especially important for learning the language of special domains, such as medicine. This is the reason for the formation of our UKRainian MEDicine text corpus – UKRMED, which is described in [3]. UKRMED was formed specifically to study the complexity of the perception of medical texts in Ukrainian. In work [3], we suggested that all texts can be divided into three categories: simple, moderate, and complex. This is due to the different perceptions of these texts. The purpose of this study is to evaluate the complexity level of texts from our UKRMED corpus using various readability metrics. This will allow us to evaluate the hypothesis that the texts in our corpus represent three groups of perception complexity.

An important feature of studying medical texts is that the simplification or explanation of such texts for ordinary people will help them to properly prepare for examination or a visit to a doctor, properly organize taking medicine, and consult a specialist in case of important symptoms. Based on

our previous studies [2, 3], we can highlight that in Ukrainian medical texts there are a lot of borrowed words, Latin terms, and special collocations. This gives reasons to study the features of Ukrainian texts in the medical domain. Figure 1 shows the basic elements for readability. In our work, we focused only on the analysis of the medical text style, namely the sentence analysis and medical lexis.
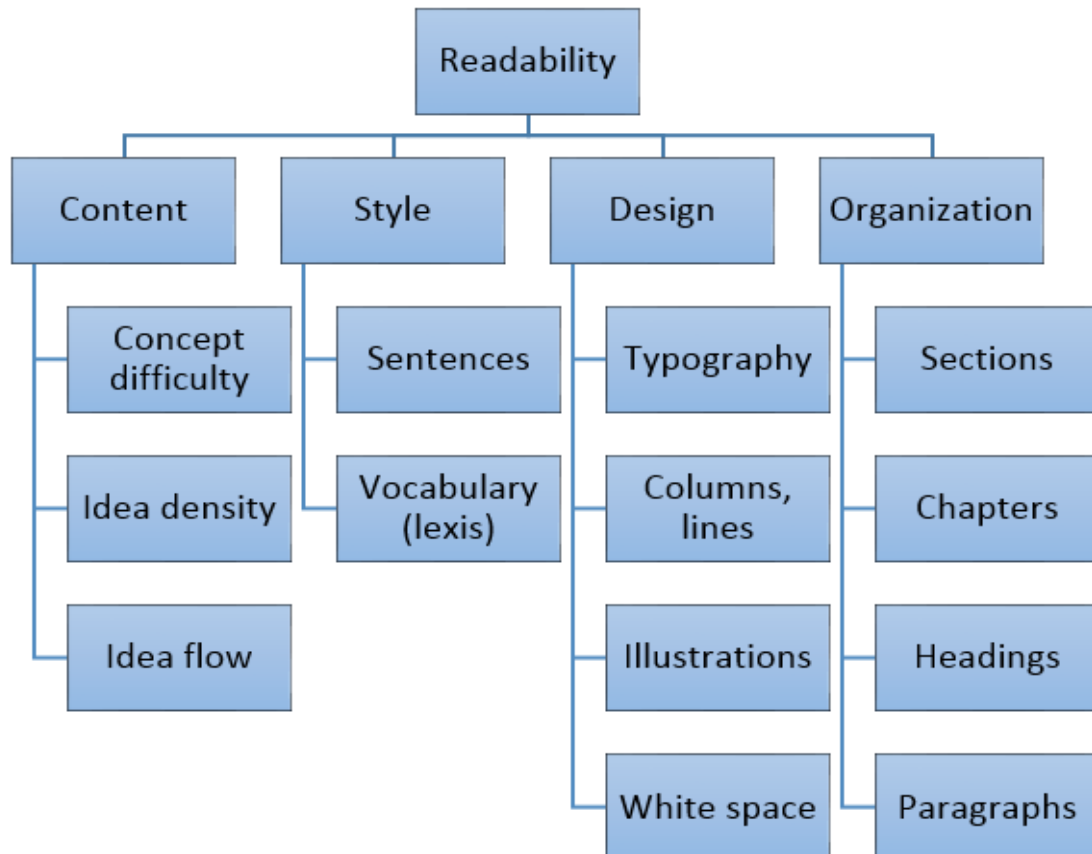
Figure 1: The main elements of readability

This research sight to (1) how the most commonly used readability formulas work with the Ukrainian texts in the medical domain, (2) compare and contrast these different formulas to various texts (blogs, protocols, and wiki texts), (3) research different medical text features which will be used for text simplification and classification medical texts and (4) prepare recommendations for using these formulas to the evaluation of readability medical texts in Ukrainian.

2. Readability formulas

Various readability indices are used to measure text complexity [1, 4, 5]. The analysis shows that the use of readability ratings allows us to assess the relevance of the text to a specific target group, to characterize the age of readers, as well as the attitude of non-native speakers to this text. When the text is too complicated or difficult to read, messages may not be understood. On the other hand, when the text is too simple, your audience may feel boring. In any case, the readability of the text affects the degree of interaction and perception of the message.

Therefore study the complexity of texts is important. Many researchers look deeply at the issue [4, 5, 6, 7]. The task of text simplification is quite wide. Paper [8] is focused on text simplification for congenitally deaf people. Authors [9] study complex-simple sentence pairs from the Newsela corpus. Newsela is the largest collection of professionally written simplifications for tasks of text simplification [9]. The complexity of the perception of questionnaires is studied by the authors [10]. In their work, they use multidimensional analytical methods.

Consumer health informatics is a field that provides health information to improve healthcare decision-making [11]. Such works as [12, 13, 14] are devoted to the evaluation of the readability of texts on medical topics.

A special place among the researches related to solving the problem of simplifying text is occupied by studies of readability metrics [15, 16, 17]. We can notice that some authors pay attention to the readability issues of medical texts [12, 14, 18]. In the study [19] a method for assessing the difficulty of words has adapted to make it more suitable to medical Swedish. In the paper [20] is underlined that poor health literacy is known to impact negatively on medical outcomes. They assess the readability of online ophthalmic literature by applying validated readability formulas: Flesch Reading Ease Score, Simple Measure of Gobbledygook, and Flesch-Kincaid Grade Level [20].

There are many formulas that measure the readability of text. Any readability formula represents the method of measuring or predicting the difficulty level of text. Following the deep literature analysis, we can highlight the most popular readability formulas.

It is a well-known issue that readability formulas are used for the evaluation of written information. However, we underline that evaluation under the readability formula results varies considerably due to the language or domain area features. These variations caused uncertainty of interpretations of reading grade level estimates. Next, we will consider the most commonly used readability evaluation methods (https://readable.com/features/readability-formulas/).

Let us consider the set of readability formula chosen for our text corpus evaluation.

The Flesch Reading Ease index [4, 17]. It is computed based on the average number of syllables per word and the average number of words per sentence (1). Nowadays, this Flesch test is one of the most widely used, most tested, and reliable readability formulas [21].

$$206.835 - 1.015 * \left(\frac{total\ words}{total\ sentences} \right) - 84.6 * \left(\frac{total\ syllables}{total\ words} \right). \tag{1}$$

Flesch-Kincaid Grade Level [21]. It computes readability based on the average number of syllables per word and the average number of words per sentence (2). The score indicates a grade-school level. The higher the reading score, the easier a piece of text is to read.

$$0.39 * \left(\frac{total\ words}{total\ sentences} \right) + 11.8 * \left(\frac{total\ syllables}{total\ words} \right) - 15.59. \tag{2}$$

Gunning's Fog Index [4, 21]. It is a weighted average of the number of words per sentence, and the number of long words per word.

The Coleman–Liau Readability Formula (Coleman–Liau index) [4, 21]. This index is calculated with the following formula:

$$CLI = 0.0588L - 0.296S - 15.8. \tag{3}$$

L is the average number of letters per 100 words. S is the average number of sentences per 100 words.

Dale–Chall readability formula [4] based on the following equation:

$$Raw\ Score = 0.1579 * (PDW) + 0.0496 * ASL, \tag{4}$$

where Raw Score – reading grade of a reader who can comprehend your text at 3rd grade or below. PDW is a percentage of difficult words and ASL – average sentence length in words.

The FORCAST readability formula [4, 21]. The formula is:

$$Grade\ level = 20 - (N / 10), \tag{5}$$

where N – number of single-syllable words in a 150-word sample.

The Automated Readability Index (ARI) [4, 21]. This index is calculated as

$$4.71 * \left(\frac{characters}{words} \right) + 0.5 * \left(\frac{words}{sentences} \right) - 21.43, \tag{6}$$

where characters are the number of letters and numbers.

We create the corpus UKRMED, the UKRainian MEDicine text corpus, with a focus on three categories of medical writing information related to their complexity [3]. Therefore, using formulas presented above, and considering their applicability to medical texts, we intend to evaluate the texts from our corpus UKRMED and confirm our assumptions about the different complexity of the collected texts.

3. Experiments with readability formulas on our corpus
3.1. Data description

The common requirement of the corpus is providing data for language issues study. The information about our data is given in Table 1.

Table 1
Category distribution and quantitative data of the UKRMED corpus

Category	Number of sentences	Number of tokens
Simple	25,395	320,209
Moderate	27,081	363,539
Complex	26,730	329,837
Total	79,209	1013,585

UKRMED is created to study medical text simplification and for experiments with readability metrics. In our previous works [2, 3], we calculate some featured indices for our text corpus. We try to collect texts under the balance, i.e. text length in tokens is 363,539 for Simple texts, 320,209 for Moderate texts, and 329,837 for Complex texts that are quite similar.

As a result of our experiments, we calculated statistical features on the lexical, syntactic, and paragraph levels. Also, we received parts of speech categorization for our three categories and analyzed them. More information about UKRMED is presented in [3] work.

3.2. Analysis different features of UKRMED corpus

Based on the analysis of variety publications to determine the complexity of the text [1, 4, 18, 21], we have identified a number of properties, the values of which we calculated for our corpus of medical texts. These properties are presented in Table 2. All properties are divided into several categories: phonological, morphological, syntactic, and inter-sentential features.

Table 2
Research features for the UKRMED medical text corpus

	Phonological Features
1	Number of syllables per word
	Morphological Features
1	Morphological diversity
2	Length of nouns per word
3	Length of adjectives per word
4	Mean length of adverbs per word
5	Lexical sophistication
6	Verb sophistication
7	Number of unique words
8	Ratio of unique words to total words (TTR)
9	Guiraud's corrected TTR (GTTR)
10	Caroll's corrected TTR (CTTR)
11	D measure
12	Measure of Textual Lexical Diversity (MTLD)

Syntactic Features	
1	Mean length of the sentence
2	2 ngrams per word
3	3 ngrams per word
4	4 ngrams per word
5	2 ngrams per sentence
6	3 ngrams per sentence
7	4 ngrams per sentence
Inter-Sentential Features	
1	Connectors, such as and, therefore, and hence, indicate long and elaborate sentences as well as an advanced structure of the text (#connectors)
2	Argumentative discourse connectors are a subset of discourse connectors that indicate a higher level of reasoning and argumentation (#argumentative_connectors)
3	Connectors sentences feature - (#connectors/ #sentances)
4	Argumentative connectors sentences - (#argumentative_connectors/ #sentances)

Indicators (features) that are shown in the Table 2 were detailed described in the work [4], but below we give a brief description of some of them. The morphological diversity is calculated as

$$\frac{number\ of\ lemmas}{total\ number\ of\ lemmas}. \tag{7}$$

The verb sophistication measure (VSM) estimates the number of sophisticated verbs in relation to the total number of verbs

$$\frac{number\ of\ sophisticated\ verbs}{number\ of\ verbs}. \tag{8}$$

Lexical sophistication reflects percentage of sophisticated or advanced words in a text. There are different definitions of sophisticated vocabulary. We consider that the word is sophisticated in case its frequency rank is over 3000.
Guiraud's corrected TTR (GTTR) is calculated as

$$GTTR = \frac{NDW}{\sqrt{n}} * 100^n. \tag{9}$$

Carroll's lexical diversity measure or Caroll's corrected type-token ratio (CTTR) is calculated as

$$CTTR = \frac{number\ of\ types}{\sqrt{2n}}. \tag{10}$$

The D measure is based on the predicted decrease of the TTR according to the size of the text. The Measure of Textual Lexical Diversity (MTLD) evaluates the lexical diversity in another way. MTLD is designed to reduce the effect of the text length. MTLD is calculated as the mean length of strings in a text that has a given TTR value.

For all these features from Table 2 we received values for each text per category of our dataset (Table 3). In Table 3 we stayed only that features that differ depending on the category of texts. The rest of the meanings are very close and did not change in any way depending on the category of texts. For example, the morphological diversity for per text category is showed on Figure 2.

The values from Table 3 are showed that the most important features are lexical sophistication, mean length of the sentence, argumentative discourse connectors, measure of textual lexical diversity and argumentative connectors sentences. Other indicators are differ, but not so much. Therefore, these are features could be used for medical text classification or text simplification.

Table 3

The feature values of UKRMED corpus

Feature	Simple texts	Moderate texts	Complex texts
Lexical sophistication (frequency rank is over 3000)	182,51	195,91	156,44
Verb sophistication (frequency rank is over 3000)	205,72	225,62	180,41
Number of unique words	222,16	241,27	216,00
Ratio of unique words to total words (TTR)	63,83	61,49	59,10
Guiraud's corrected TTR (GTTR)	1176,87	1179,41	1120,61
Caroll's corrected TTR (CTTR)	832,17	833,97	792,39
Measure of Textual Lexical Diversity (MTLD)	378,44	356,88	292,22
Mean length of the sentence	121,04	147,96	162,21
3 ngrams per word	0,62	0,54	0,54
4 ngrams per word	0,94	0,89	0,89
Argumentative discourse connectors	0,075	0,049646	0,038957
Connectors sentences feature	1,209	1,220794	1,232549
Argumentative connectors sentences	0,075	0,049646	0,038957
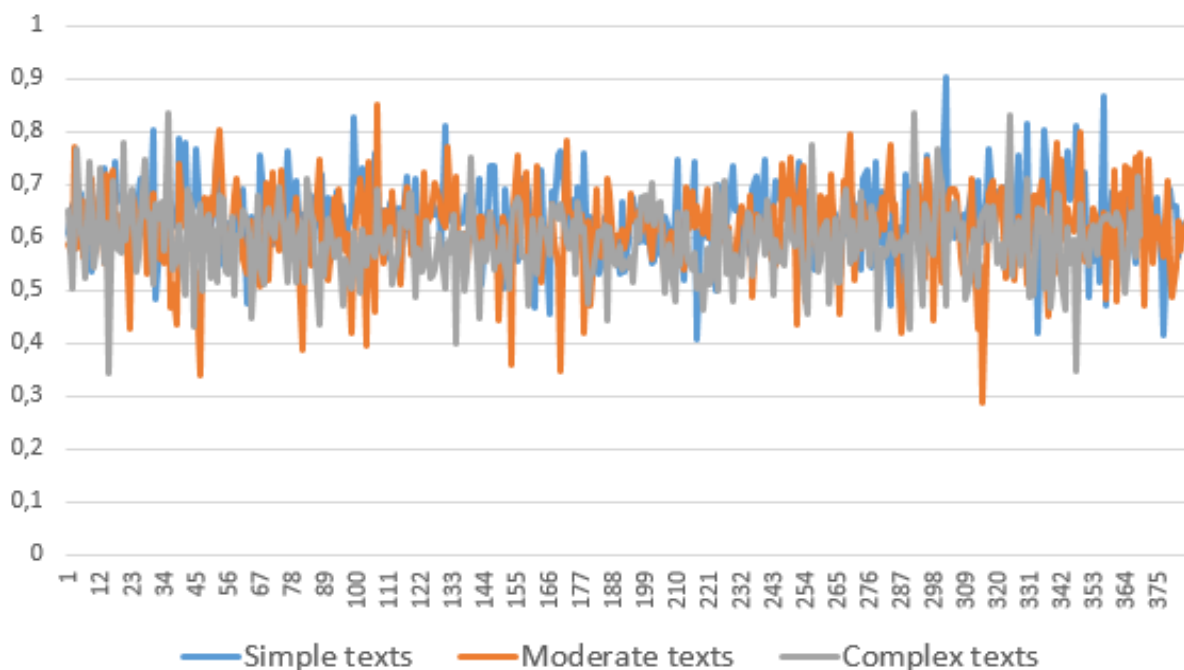
Figure 2: Morphological diversity for per text category

3.3. Experiments with readability formulas for medicine domain

In this section, we tried to analyze the received results and interpreted theirs for our data and domain. Firstly, it should be noted that we formed our corpus in a certain way, breaking it into three categories (genres). We suggested that texts from the category "*Simple texts*" will be easy to understand, as they are taken from blogs, forums, etc. and written in a lively and simple language for most readers. Texts from the category "*Complex texts*" will be the most difficult, since they represent clinical protocols, medical scientific articles, etc., but texts from the category "*Moderate texts*" will be somewhere between simple and complex, since there are also Wikipedia articles simple enough, but

sometimes complex. It depends on the article author. It should also be noted that none of the formulas was adapted for the Ukrainian language.

We calculated Gunning Fog Index, Flesch-Kincaid Grade Level, Coleman–Liau index, Dale–Chall readability formula, the FORCAST readability formula and the Automated Readability Index (ARI) for all text categories of our corpus. All values are given in Table 4.

Table 4
The feature values of UKRMED corpus

Readability Formulas	Simple texts	Moderate texts	Complex texts
Gunning Fog Index	36,63	41,82	44,62
Flesch-Kinkaid Grade Level	-134,66	-149,47	-156,02
Coleman–Liau index	43,69	47,72	49,95
Dale–Chall readability formula	8,67	9,61	10,15
The FORCAST readability formula	17,16	17,58	17,47
The Automated Readability Index (ARI)	34,19	38,01	40,20

Consider the Flesch-Kincaid Grade Level. We have obtained very low values and with a minus. It's mean that texts are very difficult for the majority of people.

As a confirmation of our results, we used the site LeStCor (http://www.lestcor.org/). This resource was created for the calculation of different readability indices for the Russian language. Because Ukrainian and Russian are kindred languages, we used this resource for our experiments too. We received the following message "Very difficult to read. Best understood by university graduates". So, all texts from all categories are very difficult. However, our hypothesis was confirmed, because we received the lowest value for "Simple texts" category, and the highest value for the "Complex texts".

The Gunning Fog Index, Coleman–Liau index, Dale–Chall readability formula and ARI have the same trend. Only The FORCAST readability formula doesn't feet the common tendency and has the highest value for the "Moderate texts".

3.4. Analysis of difficult lexica in the corpus

After we received the results of experiments on our corpus using readability formulas, we decided to mark in our corpus the elements that cause the reader the greatest difficulty in perceiving and understanding the text. We asked volunteers (master students) to labeling words, phrases and sentences in the texts for understanding.

We have not yet managed to process all the texts in our corpus, but for the first experiments, we received 140 texts – moderate category, 143 texts – simple, and 148 texts – moderate. As an analysis results of the marked elements in these documents, we received the following information, is presented in Table 5.

Table 5
Statistic of difficult elements in the UKRMED

Text category	Difficult words	Difficult phrases	Difficult sentences
Simple texts	314	229	25
Moderate texts	601	357	34
Complex texts	802	372	57

After we removed the duplicate words and phrases, the number of words decreased, but still there are quite a lot of them (Table 6).

Table 6

Statistic of difficult elements in the UKRMED (after duplicate removing)

Text category	Difficult words	Difficult phrases	Difficult sentences
Simple texts	217	229	25
Moderate texts	420	357	34
Complex texts	429	372	57

A detail analysis of Tables 5 and 6, you can see that the category of texts "Simple texts" is really the easiest to understand, it has the least complex words, phrases and sentences. The most difficult category of texts is "Complex texts", but after reducing duplicates, the category "Moderate texts" is closer to "Complex texts". The result of the labeling showed that our assumptions that we used when forming this corpus were confirmed. The gradation of the text categories is correct.

In this work, we decided to focus not on all the complex elements that were involved in the markup, but only on words. Since, firstly, they prevail in all categories and cause more difficulty for the reader, and secondly, they are then actively used to simplify the text.

Here are examples of complex words that were highlighted during the labeling (Table 7).

Table 7 shows Top-10 difficult words that caused the reader to understand difficulties. When analyzing all the words, we identified three categories of words that are complex: 1 – abbreviations, 2 – medical special terms, 3 – noise words, words that were mistakenly and are commonly used words.

If we consider these words from the point of view of their further use in the process of simplifying the text, then both abbreviations and medical special terms can be explained using available definitions, external dictionaries of medical vocabulary, and other linguistic resources.

We decided to see if complex words are found in phrases and sentences. Perhaps it is they that cause the reader's difficulty in perceiving of the phrase or a whole sentence. To do this, we checked how often complex words occur in a particular category in phrases and sentences.

Table 7

Top-10 difficult words in the UKRMED (for all categories)

Complex texts	Simple texts	Moderate texts
Гіпоглікемія/Hypoglycemia	Коагулограма/Coagulogram	Інвагінації/Intussusception
Ектопія/Ectopia	Телофаза/Telophase	Аміодарон/Amiodarone
КВГ	РМЗ	Атріовентрикулярна/Atrioventricular
ЛПМД	Біопсія/Biopsy	Гебефре́нія/Hebephrenia
Х-хроматин/X-chromatin	Гіперхромність/ Hyperchromia	Езофагіт/Esophagitis
Аміодарон/Amiodarone	Кокоглицинат/Cocoglycinate	ЛПНЩ
Гемокультура/Blood culture	Меланоми/Melanoma	ПМЦ
Гіпоспадія/Hypospadias	Пневмокок/Pneumococcus	Римантадин/Rimantadine
Кармін/Carmine	Синапси/Synapses	Адгезіоліз/Adhesiolysis
Патогенез/Pathogenesis	Цитокінез/Cytokinesis	Амітриптилін/Amitriptyline

Table 8 showed that complex words are found in phrases more often than in sentences, but at the same time, in ratio to the total number of phrases and sentences for each category of texts, this is a fairly small percentage. The number of compound words that were found in phrases and sentences to the total number of compound words is no more than 0.06%. Therefore, we can conclude that individual complex words (abbreviations, special medical terms, etc.) do not have a large impact on the complexity of the phrase and sentence.

Consider an example of a sentence that is difficult to understand:

*«Зазвичай уражаються метастазами последовательнокаждая група, але нерідко бувають винятки і метастази можуть бути знайдені в проміжній або базальної групі, а епіпараколіческіе лімфатіческіеузли залишаються інтактними.По топографії лімфометастазов раку слепой висхідної ободової кишки для радикального видалення зон регіонарного метастазірованіянеобхідма правобічна геміколектомія з **резекцією**…».*

("Usually affected by metastases sequentially each group, but often there are exceptions and metastases can be found in the intermediate or basal group, and epiparakolicheskie lymph nodes remain intact. According to the topography of lymphometastases of cancer of the cecum and ascending colon for radical removal of areas of regional metastasis, a right-sided hemicolectomy with resection is required.")

In this sentence, the complex word «*резекцією*» («*resection*») was found, but if we look at the whole sentence, we understand that it, in principle, contains many other compound words, long enough and difficult for a person who is not a specialist in medicine. And we don't have these other difficult words in the list of difficult words. Perhaps this is due to the fact that the evaluators did not correctly mark up the texts, or perhaps because this sentence is so incomprehensible and complex that the evaluators decided to place it in complex sentences, but not to single out individual complex words in it.

Table 8
The frequency of complex words occurrence in phrases and sentences

Text category	Difficult phrases	Difficult sentences
Simple texts	4	0
Moderate texts	10	2
Complex texts	20	1

4. Discussion and future work

Linguistically complex tasks, such as the medical text understanding, are the most challenging because they require linguistic intuition. This task is a rather complicated and depends on many factors such as language, subject area, etc. For the Ukrainian language, this direction is only beginning to develop, and therefore there are no large results in this area. Our experiments showed that the use of readability formulas could help us in this task, and we must look for other methods to test the complexity of the medical text.

Quality of text corpora is the key to obtaining good results. The problem is that there is a lack of texts to form corps in specific areas, such as medicine. The Ukrainian language is also in the early stages of research. Our corpus UKRMED still has many shortcomings, but it is the first step towards solving the problem of simplifying the Ukrainian medical text.

We collected the texts under the assumption that three categories of text complexity can be distinguished. The main idea of our research is the simplification of the medical text depends on the complexity of this text and the stakeholder, who studies this text. Therefore, we evaluated all texts by readability metrics.

In our work, we analyzed the most commonly used readability formulas in health care literature. Readability estimates using readability formulas were compared for different genres in the medicine domain. We apply our own perception and attitude of medical texts to divide them into three categories. Readability formulas demonstrated sometimes very similar results, but sometimes not. However, all texts are very difficult for understanding in general meaning.

In future, we plan to mark our corpus as follows. Each document will have three types of markup, the first type will mark out complex lexis (medical terms), the second – sentences that are difficult to understand, and the third type will contain a text complexity label (easy, intermediate, and difficult). Such markup will allow a qualitative classification of our texts, as well as preparatory work to identify the complex elements of a medical text for its further simplification.

5. Acknowledgements

6. References

[1] S. Vajjala, D. Meurers, Readability assessment for text simplification: From analysing documents to identifying sentential simplifications, ITL – International Journal of Applied Linguistics 165(2) (2014) 194–222. doi:10.1075/itl.165.2.04vaj.

[2] O. Cherednichenko, O. Kanishcheva, N. Babkova, Complex term identification for Ukrainian medical texts, Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine (IDDM 2018), Vol. 2255, 2018, pp. 146–154.

[3] O. Cherednichenko, O. Kanishcheva, O. Yakovleva, D. Arkatov, Collection and Processing of a Medical Corpus in Ukrainian. Proceedings of the 4 Int. Conf. On Computational Linguistics and Intelligent Systems (COLINS), volume I: Main Conference CEUR-WS. Vol. 2604, 2020, pp. 272–282.

[4] M. Z. Kurdi, Text Complexity Classification Based on Linguistic Information: Application to Intelligent Tutoring of ESL, Journal of Data Mining and Digital Humanities 2020 (2020).

[5] H. Saggion, Automatic Text Simplification. Synthesis Lectures on Human Language Technologies. 10(1) 2017 1–137. doi:10.2200/S00700ED1V01Y201602HLT032.

[6] C. Scarton, G. H. Paetzold, L. Specia, Text simplification from professionally produced corpora. Proceedings of the LREC 2018 – 11th International Conference on Language Resources and Evaluation, European Language Resources Association (ELRA), 2019, pp. 3504–3510.

[7] D. Ferrés, M. Marimon, H. Saggion, A. AbuRa'ed, YATS: Yet another text simplifier, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, Vol. 9612, 2016, pp. 335–342. doi:10.1007/978-3-319-41754-7_32.

[8] K. Inui, A. Fujita, T. Takahashi, R. Iida, T. Iwakura, Text simplification for reading assistance. Association for Computational Linguistics (ACL), 2003, pp. 9–16. doi:10.3115/1118984.1118986.

[9] C. Scarton, G. H. Paetzold, L. Specia, Text simplification from professionally produced corpora. Proceedings of the LREC 2018 – 11th International Conference on Language Resources and Evaluation European Language Resources Association (ELRA), 2019, pp. 3504–3510.

[10] S. C. Peter, J. P. Whelan, R. A. Pfund, A. W. Meyers, A text comprehension approach to questionnaire readability: An example using gambling disorder measures. Psychological Assessment 30(12) (2018) 1567–1580. doi:10.1037/pas0000610.

[11] D. Flaherty, L. Hoffman-Goetz, J. F. Arocha, What is consumer health informatics? A systematic review of published definitions. Informatics for Health and Social Care 40(2) (2015) 91–112. doi:10.3109/17538157.2014.907804.

[12] S. Alotaibi, M. Alyahya, H. Al-Khalifa, S. Alageel, N. Abanmy, Readability of Arabic Medicine Information Leaflets: A Machine Learning Approach. In Procedia Computer Science, Elsevier B.V., Vol. 82, 2016, pp. 122–126. doi:10.1016/j.procs.2016.04.017.

[13] P. Mukherjee, G. Leroy, D. Kauchak, S. Rajanarayanan, D. Y. Romero Diaz, N. P. Yuan, S. Colina, NegAIT: A new parser for medical text simplification using morphological, sentential and double negation. Journal of Biomedical Informatics 69 (2017) 55–62. doi:10.1016/j.jbi.2017.03.014.

[14] D. Kauchak, G. Leroy, Moving beyond readability metrics for health-related text simplification. IT Professional 18(3) (2016) 45–51. doi:10.1109/MITP.2016.50.

[15] S. Crossley, D. Allen, D. McNamara, Text readability and intuitive simplification: A comparison of readability formulas. Reading in a foreign language 23(1) (2011) 84–101.

[16] S. Štajner, R. Evans, C. Orăsan, R. Mitkov, What Can Readability Measures Really Tell Us About Text Complexity? Workshop on Natural Language Processing for Improving Textual Accessibility (NLP4ITA), 2012, pp. 14–21.

[17] M. Cha, Y. Gwon, H. T. Kung, Language modeling by clustering with word embeddings for text readability assessment. Proceedings of the International Conference on Information and Knowledge Management, Association for Computing Machinery, Vol. Part F131841, 2017, pp. 2003–2006. doi:10.1145/3132847.3133104.

[18] G. Leroy, S. Helmreich, J. R. Cowie, The influence of text characteristics on perceived and actual difficulty of health information. International Journal of Medical Informatics 79(6) (2010) 438–449. doi:10.1016/j.ijmedinf.2010.02.002.

[19] E. Abrahamsson, T. Forni, M. Skeppstedt, M. Kvist, Medical text simplification using synonym replacement: Adapting assessment of word difficulty to a compounding language, Association for Computational Linguistics (ACL), 2015, pp. 57–65. doi:10.3115/v1/w14-1207.

[20] M. R. Edmunds, R. J. Barry, A. K. Denniston, Readability assessment of online ophthalmic patient information. JAMA Ophthalmology 131(12) (2013) 1610–1616. doi:10.1001/jamaophthalmol.2013.5521.