Coh-metrix readability formulas for an academic text analysis

Aleksander Kiselnikov¹[0000-0003-2639-678X], Diliara Vakhitova¹[0000-0002-8851-5521], Tatiana Kazymova¹[0000-0002-7853-0461]

¹Kazan State University of Architecture and Engineering, Kazan, 420043, Russia
²Kazan (Volga region) Federal University, Kazan, 420000, Russia
E-mail: kalejandr@gmail.com

Abstract Readability formulas play a significant role in the process of text design for academic purposes. Nowadays formulas are digitalized that makes their implementation easy and convenient. Formulas are designed for different languages and are aimed at dealing with texts of different genres. In the article, we evaluate readability of academic texts, which are in the textbook “English for Engineers”. This book is recommended by the Ministry of Science and Higher Education of the Russian Federation for all technical Universities. Readability indices evaluation is performed with online tool Coh-Metrix. The tool suggests three readability indices obtained by means of three formulas: Flesch-Kincaid Grade Score, Flesch Reading Ease and Coh-Metrix L2 Reading Index. The Pearson correlation clearly demonstrates the absence of readability indices decrease from the first text in the book, which is considered to be the most readable to the last one that should be the less readable. By the way, Pearson correlation illustrates strong correlation between Flesch-Kincaid and Coh-Metrix L2 readability indices.

Keywords: Russian education, text readability, Coh-Metrix, Flesch-Kincaid Grade Score, Flesch Reading Ease, Coh-Metrix L2 Reading Index.

1 Introduction
Since 1992 the system of the Russian education has been being transformed [1]. Digitalization of educational process is a core task of modern system of education in the Russian Federation. New standards are aimed at developing new digital tools helping students acquiring information and enlarging knowledge. Among the tools there are online sources with huge database (books, lectures, presentations, videos etc.) on almost all subjects. However, this is only one way of digitalization implementation for academic purpose. Another way is the use of up-to-date software and internet tools to perform the analysis of different sources of information to evaluate the best possible solution for information presentation and its further comprehension.

We are focused on English as second language in the Russian technical University. We study the texts for reading which are in the textbook “English for Engineers” by Polyakova et. al [2]. The idea is to examine the distribution of texts throughout the book: texts are expected to be sorted according to their readability decreasing from the beginning of the book until its end. The content of each unit in the textbook is standardized: each unit provides four texts for reading labeled with letters A, B, C and D. To provide the experiment accuracy, we analyze only the texts labeled with letter “B”.

Different automated tools are able to proceed texts on different levels. Macro level is for the calculation of quantitative parameters, like syllables, words, sentences and paragraphs within a text, whereas micro level’s aim is qualitative parameters, like semantic interpretation, text cohesion and coherence evaluation, etc.

The tool for text complexity and readability analysis can obviously become an indispensable part of textbooks’ content improvement or reorganization at the stage of design.

In this article we operate with computerized automated tool Coh-Metrix to evaluate readability indices of the texts.

In this research we are guided by the two key questions:
1. What indices of readability do the texts have?
2. Are the texts organized in the book in accordance with their readability indices?
2 Literature review
Engaging with a foreign language text in the context of academic discourse students substantially encounter reading tasks. The text comes as a source of information and language knowledge. To make sure that a text can be comprehended by a student, it should be properly designed. Comprehension may be blocked by factors, like a text’s structure or vocabulary. Readability formulas are the first and simplest tool to predict a possibility of a text acquisition by a target reader. As follows, text readability should be of proper level to help acquiring information, not to inhibit it.

Nowadays many formulas are called to deal with texts of different genres in different languages. Many automated tools for text complexity evaluation, like Coh-Metrix [3] and ReaderBench [4] include readability formulas. The most used are Flesch Reading Ease (FI) and Flesch-Kincaid Grade level (FK) formulas. Modern trends determine the necessity of the formulas adaptation for texts not only in the English language but in others as well. The Russian language is not an exception. The formulas were adopted to evaluate the readability of Russian academic texts [5, 6].

In the article, we perform analysis basing on Coh-Metrix. It is a widely used tool for text analysis [7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 17; 18; 19; 20; 21; 22]. This tool provides 106 indices of text complexity. All indices reported by Coh-Metrix are normalised for text length [23]. All indices are grouped into eleven parts:
1. Descriptive;
2. Text Easability Principal Component Scores;
3. Referential Cohesion;
4. LSA;
5. Lexical Diversity;
6. Connectives;
7. Situation Model;
8. Syntactic Complexity;
9. Syntactic Pattern Density;
10. Word Information;
11. Readability [24].

In the article, we are interested primarily in the eleventh part – Readability. Here are three different formulas: Flesch Reading Ease, Flesch-Kincaid Grade Level and Coh-Metrix L2 Reading Ease (CML2).
- Flesch Reading Ease = 206.835 – (1.015 x number of words/number of sentences) – (84.600 x number of syllables/number of words);
- Flesch-Kincaid Grade Level = (0.39 x number of words/number of sentences) + (11.8 x number of syllables/number of words) – 15.59;
- Coh-Metrix L2 Reading Index = – 45.032 + (52.230 x Content Word Overlap Value) + (61.306 x Sentence Syntax Similarity Value) + (22.205 x CELEX Frequency Value) [3].

Flesch Reading Ease and Flesch-Kincaid Grade level are to provide indices of readability that correlate a text to a particular academic age of L1 reader. However, “readability formulas are less predictive than models of text comprehension, processing, and familiarity derived from advanced natural language processing tools” [25] they are still widely used.

Coh-Metrix L2 Readability formula is supposed to reflect cognitive and psycholinguistic processes of reading. It deals with three variables: CELEX Word Frequency (logarithm mean for content words), Sentence Syntax Similarity (sentence to sentence adjacent mean), and Content Word Overlap (proportional adjacent sentences unweighted) [26]. The operation with such variables significantly differs this formula from traditional readability formulas, which, primarily, take into consideration just formal aspects of a text, like an average length of a word or an average length of a sentence but without any respect to the content of a text. In this sense, Coh-Metrix L2 Readability formula relates more to simple text complexity evaluation tools whose aim is to describe a text in accordance with its syntax, vocabulary, morphology, semantics, etc. Another reasonable fact to implement this formula for our texts analysis is that it is designed based on a corpus of academic texts.
Materials and methods
The evaluation of readability indices was carried out on thirteen academic texts taken from the textbook “English for Engineers” employed by a technical university students learning English as a second language. The texts are labeled T1–T13 respectively.

Stage 1. Each text was successively analyzed on the automated computational platform Coh-Metrix (version 3.0). Out of 106 Coh-Metrix indices only Readability indices were selected.

Stage 2. We analyzed how the texts are located in the book: whether readability indices demonstrate the gradual decrease.

Stage 3. As a final step, we implemented Pearson correlation coefficient formula to compute the strength and direction of a linear relationship between:

1) Flesch-Kincaid Grade level index and a text position in a book;
2) Flesch Reading Ease index and a text position in a book;
3) Coh-Metrix L2 Reading index and a text position in a book;
4) Flesch-Kincaid Grade level index and Flesch Reading Ease index;
5) Flesch-Kincaid Grade level index and Coh-Metrix L2 Reading index;
6) Flesch Reading Ease index and Coh-Metrix L2 Reading index.

Results
As the experiment supposes, we analyze texts from each unit of the textbook, which are parts of the “Reading”. Originally, each unit contains four texts for reading labeled with letters A, B, C, D. To ensure the accuracy of the experiment we take only texts labeled with letter B. Thus, we have thirteen texts for analysis. As we have texts located gradually in each unit since the beginning of the book, we expect them to decrease their readability.

Figure 1 represents FK Grade level indices; Figure 2 – Flesch Reading Ease indices; Figure 3 – Coh-Metrix L2 Reading indices of thirteen texts (T1–T13) in accordance with the automated computational platform Coh-Metrix.
Figure 3 demonstrates the range of FK readability indices for 13 analyzed texts. The higher the index is, the less readable the text is. The borders of indices are within 4.0 and 10.6. It refers all the texts to the school students from the 4th till the 10th grade. However, we do not notice any decrease of texts’ readability indices since the beginning of the book until its end. Texts demonstrate random distribution of readability indices.

Figure 2 demonstrates the range of Flesch Reading Ease readability indices for 13 analyzed texts. The higher the index is the more readable the text is. The borders of indices are within 54.2 and 82.1. It refers all texts to the school students from the 6th grade till the high school students. No decrease of a text readability is evaluated.

Figure 3, that represents the results of Coh-Metrix L2 Reading indices analysis for 13 texts, also does not demonstrate any decrease of readability.

5 Discussion
To support the idea that the remoteness of a text from the beginning of the book, in the case study, does not influence its readability we applied Pearson correlation.

Pearson correlation coefficient \( r \) between texts’ remoteness from the beginning of the book and their FK indices is \( r=0.135899 \). The absence of correlation is obvious.

Pearson correlation coefficient \( r \) between texts’ remoteness from the beginning of the book and their FI indices is \( r=-0.22377 \). It demonstrates the absence of correlation.

Pearson correlation coefficient \( r \) between texts’ remoteness from the beginning of the book and their CML2 indices is \( r=0.110943 \). Thus, the correlation is not detected.

Although, thirteen texts labeled with letter B in the textbook “English for Engineers” are distributed in the book without any respect to readability indices it seems interesting to examine whether there any correlation between FK, FI and CML2 readability indices is.

The interpretation of FK indices demands clarification. In terms of the formula, the lower the index is, the more readable the text is and vice versa. The FI and CML2 readability formulas suppose another principle of work – the higher the obtained indices, the more readable the texts are.

First, Pearson correlation between indices obtained by FK and FI formulas which is \( r = -0.88915 \). However, we see a strong negative effect of FK indices on FI indices, we should converse the result into the positive effect. This correction is necessary due to the different approach to readability level demonstration in the formulas themselves.

Second, we calculated Pearson correlation between FK indices on CML2 indices. We detected an obvious negative effect of FK Grade level indices on CML2 indices which is \( r = -0.71131 \). However, the result should be conversed like in the previous calculation.

Third, Pearson correlation between Flesch Reading Ease indices and Coh-Metrix L2 indices is \( r = 0.625946 \). The correlation is rather obvious but less strong than between FK and CML2 indices.

6 Conclusions and future work
The research demonstrated the implementation of Coh-Metrix tool to evaluate readability of texts, which are in the textbook “English for Engineers”. Coh-Metrix provided 106 indices, but we
concentrated on the three of them characterizing readability: Flesch Reading Ease, Flesch-Kincaid Grade Level and Coh-Metrix L2 Reading Index.

The analysis clearly illustrated no implementation of readability formulas for the texts’ choice/design for the textbook under the study as the texts are distributed throughout the book without decrease of readability as it was expected to be.

What is interesting, Flesch Reading Ease and Flesch-Kincaid Grade level indices, based on quantitative parameters, have a strong Pearson correlation with Coh-Metrix L2 Reading indices, which are based on lexis and syntax of a text.

Future work supposes the analysis of the texts labeled with letters A, C and B in the same book. Moreover, we plan to perform the readability analysis of other books for technical universities, recommended by the Ministry of Science and Higher Education of the Russian Federation.

Acknowledgements
The research was supported by the subsidy of the Russian Government to support the Program of Competitive Growth of Kazan Federal University.

The research presented in Parts 1 and 2 of the article was supported by the subsidy of the Russian Government to support the Program of Competitive Growth of Kazan Federal University.

The research presented in Parts 3–5 of the article was financially supported by the Russian Science Foundation, grant № 18-18-00436.

References
[1] Solnyshkina M I, Solovova E N, Harkova E V and Kiselevkov A S 2016 Language Assessment course: Structure, Delivery and Learning Outcomes International Journal of Environmental & Science Education 11(6) pp 1223–1229 DOI: 10.12973/ijese.2016.392a
[2] Polyakova T Yu, Sinyavskaya E V, Tynkova O I and Ulanovskaya E S 2016 Anglijskij yazyk dlya inzhenerov=English for Engineers (Moscow: Izdatel'skij centr «Akademiya»)
[3] McNamara D S, Graesser A C, McCarthy P M and Cai Z 2014 Automated evaluation of text and discourse with Coh-Metrix (Cambridge: Cambridge University Press)
[4] Dascalu M, Gutu G, Ruseti S, Paraschiv I C, Dessus P, McNamara D S, Crossley S A, and Trausan-Matu S 2017 ReaderBench: A Multi-lingual Framework for Analyzing Text Complexity EC-TEL pp 495–499
[5] Ivanov V V, Solnyshkina M I and Solovyev V D 2018 Efficiency of text readability features in Russian academic texts Komp'yuternaja Lingvistika i Intellektual'nye Tehnologii 2018(17) pp 267-283
[6] Solovyev V, Solnyshkina M, Ivanov V and Batyrshin I 2019 Prediction of reading difficulty in Russian academic texts Journal of Intelligent & Fuzzy Systems 36(5) pp 4553-63
[7] Aryadoust V and Liu S 2015 Predicting EFL writing ability from levels of mental representation measured by Coh-Metrix: A structural equation modeling study Assessing Writing 24 pp 35–58 DOI:10.1016/j.asw.2015.03.001
[8] Aryadoust V 2016 Understanding the growth of ESL paragraph writing skills and its relationships with linguistic features Educational Psychology 36(10) pp 1742–1770 DOI: 10.1080/01443410.2014.950946
[9] Chang Y and Stone D 2019 Why does decomposed audit proposal readability differ by audit firm size? A Coh-Metrix approach Managerial Auditing Journal 34(8) pp 895–923 DOI: 10.1108/MAJ-02-2018-1789
[10] Crossley S A and Kyle K 2018 Assessing writing with the tool for the automatic analysis of lexical sophistication (TAALES) Assessing Writing 38 pp 46–50 DOI: 10.1016/j.asw.2018.06.004
[11] Elgort I 2017 Blog posts and traditional assignments by first- and second-language writers Language Learning and Technology 21(2) pp 52–72
[12] Kalantari R and Javad G 2017 Lexical Complexity Development from Dynamic Systems Theory Perspective: Lexical Density, Diversity, and Sophistication International Journal of Instruction 10(4) pp 1–18 DOI: 10.12973/ijj.2017.1041a

[13] Lu X 2017 Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment Language Testing 34(4) pp 493–511 DOI: 10.1177/0265532217710675

[14] MacArthur C, Jennings A and Philippakos Z 2019 Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? Reading and Writing 32(6) pp 1553–74 DOI: 10.1007/s11145-018-9853-6

[15] McNamara D, Crossley S, Roscoe R, Allen L and Dai J A hierarchical classification approach to automated essay scoring Assessing Writing 23 pp 35–59 DOI: 10.1016/j.asw.2014.09.002

[16] Medimorec S, Pavlik P, Olney A, Graesser A and Risko E 2015 The language of instruction: Compensating for challenge in lectures Journal of Educational Psychology 107(4) pp 971–990 DOI: 10.1037/edu0000024

[17] Mercer S, Keller-Margulis M, Faith E, Reid E and Ochs S 2019 The Potential for Automated Text Evaluation to Improve the Technical Adequacy of Written Expression Curriculum-Based Measurement. Learning Disability Quarterly 42(2) pp 117–128 DOI: 10.1177/0731947818803296

[18] Perin D and Lauterbach M 2018 Assessing Text-Based Writing of Low-Skilled College Students International Journal of Artificial Intelligence in Education 28(1) pp 56–78 DOI: 10.1007/s40593-016-0122-z

[19] Polio C and Yoon H 2018 The reliability and validity of automated tools for examining variation in syntactic complexity across genres International Journal of Applied Linguistics (United Kingdom) 28(1) pp 165–188 DOI: 10.1111/ijal.12200

[20] Solnyshkina M I and Kiselnikov A S 2015 Text complexity: Study phases in Russian linguistics Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya 6(38) pp 86–99 DOI: 10.17223/19986645/38/7

[21] Wiley J, Hastings P, Blaum D, Jaeger A, Hughes S, Wallace P, Griffin T and Britt M Different Approaches to Assessing the Quality of Explanations Following a Multiple-Document Inquiry Activity in Science International Journal of Artificial Intelligence in Education 27(4) pp 758–790 DOI: 10.1007/s40593-017-0138-z

[22] Joshua W, Rod R and Yusra A 2017 Automated formative writing assessment using a levels of language framework Assessing Writing 34 pp 16–36 DOI: 10.1016/j.asw.2017.08.002

[23] Crossley S A and McNamara D S 2012 Predicting Second Language Writing Proficiency: The Roles of Cohesion and Linguistic Sophistication Journal of Research in Reading 35(2) DOI: 10.1111/j.1467-9817.2010.01449.x

[24] Coh-Metrix Homepage http://www.cohmetrix.com (last accessed 2020/03/17)

[25] Crossley S, Skalicky S, Dascalu M, McNamara D and Kyle K Predicting Text Comprehension, Processing, and Familiarity in Adult Readers: New Approaches to Readability Formulas Discourse Processes 54(5–6) pp 340–359 DOI: 10.1080/0163853X.2017.1296264

[26] Crossley A and McNamara D 2011 Text readability and intuitive simplification. A comparison of readability formulas Reading in a Foreign Language 23(1) pp 84–101