

A Method for Assessment of Text Complexity Based on Knowledge Graphs

Vladimir Ivanov^a, Marina Solnyshkina^b

^a*Innopolis University, Innopolis, Russian Federation*

^b*Kazan Federal University, Kazan, Russian Federation*

Abstract

The study explores the problem of assessing text complexity. In this paper we focus on measuring conceptual complexity and propose using knowledge graphs to this end. On the first stage of the research, RuThes-Lite thesaurus, a linguistic knowledge base with a total size of over 100,000 text entries (words and collocations), was used to elicit concepts in the texts of schoolbooks and represent text fragments as graphs. In the second series of experiments, we assessed complexity of English texts using knowledge graphs WordNet and Wikidata. Finally, we identified graph-based semantic characteristics of texts impacting complexity. The most significant research findings include identification of statistically significant correlations of the selected features, such as node degree, number of connected nodes, average shortest path, with text complexity.

Keywords

text complexity, thesaurus, knowledge graphs


1. Introduction


Of the three generally accepted levels of text complexity, i.e., lexical, syntactic and semantic/informational/conceptual, the third one is evidently more intricate to scrutinize and is universally recognized as the least explored [1]. It is the semantic level of a text, defined as the amount of background knowledge required to comprehend a text, that to a great extent facilitates text comprehension. Automatic measurement of lexical and syntactic complexity of Russian texts has been proposed in a number of studies [2, 3, 4] and it is not unexpected as these two levels are easier to formalize. The predominant approaches in previous work on Russian text complexity combine lexical and syntactic features [5]. As for the influence of the semantic level on text complexity, the studies conducted are still few and mostly for English.


In this article, we are developing a new approach to defining the conceptual complexity of texts through knowledge databases such as WordNet. The conceptual complexity of the text is viewed as the amount of knowledge (in particular, from the thesaurus) necessary for understanding a text. Comprehension of conceptually complexity texts requires substantial background knowledge as well as knowledge of abstract notions.


Evaluation of the proposed approach implies using a set of texts of different conceptual complexity. One possible way to solve the problem is to abridge original texts and use the

Proceedings of the Linguistic Forum 2020: Language and Artificial Intelligence, November 12-14, 2020, Moscow, Russia

 nomemm@gmail.com (V. Ivanov); mesoln@yandex.ru (M. Solnyshkina)

 0000-0003-3289-8188 (V. Ivanov); 0000-0003-1885-3039 (M. Solnyshkina)

 © 2020 Copyright for this paper by its authors.

 Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

abridged versions as part of the Corpus. This idea has been implemented in [6]. However, it seems important to assess the method on authentic, not artificial texts. A natural example of this kind of a set of texts of different conceptual (not just formal, lexical and syntactic) complexity are school textbooks of different grades which we use as the material in the current study.

An adequate conceptual level of a text complexity is especially important for readers with an insufficient level of knowledge [7, 8], in particular, for schoolchildren. In the situations when schoolchildren experience a lack of necessary knowledge, it may lead to difficulties in text comprehension. Thus, while developing educational materials for a specific audience, a learning material designer is expected to be aware of the approximated amount of theoretical and practical knowledge which the target reading audience can employ. For this purpose, we implement a corpus of school textbooks as a reference corpus.

2. Related works

2.1. Analysis of conceptual text complexity

A deeper level of semantic analysis, also referred to as conceptual analysis [6], implies taking into account semantic and pragmatic links between concepts in a text. Although it is the conceptual level that presents a real complexity of a text, so far, feasibility and methods of measuring the conceptual level of text complexity have remained unexplored.

The notion of the conceptual complexity of a text is viewed as related to the number of abstract concepts verbalized in a text, i.e., abstract words incidence, in a text [1]. The correlation between text abstractness and “linguistic complexity” was convincingly proved in [9] where the author used Russian texts as the material for the study. In a similar study, A. Laposhina [10] found out that only one of the four groups of text lexis which she identified with the help of ABBYY COMPARENO, i.e., words denoting abstract concepts, could be used as an indicator of text complexity. The groups A. Laposhina classified included the following: (1) ‘lex_physical’, i.e. nouns denoting specific material objects, including people (e.g. ‘cutlet’, ‘table’, ‘mom’); (2) ‘lex_virtual’, i.e. virtual, intangible objects, e.g. ‘base’, ‘internet’; (3) ‘lex_abstract’, words denoting abstract concepts including terms (e.g. ‘avantgarde’, ‘whim’, ‘effacement’), and (4) ‘lex_substance’, names of substances, e.g. ‘silver’, ‘vinegar’. Elements of the semantic approach, similar to that in the afore-cited paper, are implemented in [11] in which the authors apply latent semantic analysis to determine semantic proximity of text fragments. However, in most studies published, researchers analyze not a corpus or a text as a whole, but adjacent sentences or paragraphs only.

An important component of text complexity is also text cohesion that has been investigated in a number of studies [12, 11]. In [11], the notion of cohesion is defined as a concept uniting referential cohesion and deep cohesion. The first indicates how concepts in a sentence or adjacent sentences overlap and is manifested in repeated words, stems, arguments, etc. The second establishes cohesion due to the sequence of tenses, use of subordinate and connecting conjunctions and other means. In [13], cohesion is viewed as a notion verbalized with lexical chains. The latter refer to a number of adjacent words with similar meanings. Thornbury (2005) emphasizes the importance of lexical chains for maintaining text cohesion arguing that “*The*

lexical connectors include repetition and the lexical chaining of words that share similar meaning” [13]. The author also provides an example of a set of isolated sentences that have matching grammatical categories, but do not form a coherent text: “*The university has got a park. It has got a modern tram system. He has got a swimming pool.*”

As for semantic similarity of words, there are two different approaches to quantifying the extent to which words have similar meanings. The first approach (Latent Semantic Analysis) uses statistical characteristics of words in a text: frequency, co-occurrence. It was successfully used in a number of papers on text complexity, however, as noted above, mainly for local analysis.

Another possible approach is to explore semantic relations between words in a text. To this end, a researcher can use the information presented in semantic networks. The theory of semantic networks began developing half a century ago [14] with the purpose of explaining the structure of human memory. By present there have been developed numerous models of networks and there is an extensive bibliography on the topic.

The research performed in [15] indicates that many of the semantic networks studied, including those built on the basis of psycho-semantic associative experiments, have similar statistical characteristics.

2.2. Knowledge Graphs

In modern studies aimed at word processing, the two most popular and more frequently used resources are thesauri (or lexical ontologies), such as WordNet [16], and Wikidata¹, i.e. a knowledge graph with multiple semantic relationships between concepts. WordNet was originally created to study human memory which enable to expect that the use of structures like WordNet will make it possible to advance in understanding the problem of text complexity.

The semantic proximity of words is determined by their closeness in the structure of the knowledge base or thesaurus. Thesauri, however, are also one of the types of knowledge bases, since, unlike traditional dictionaries, they contain not only linguistic but also extralinguistic information, i.e. world knowledge. The latter is registered in thesauri with hypo-hyperonymic connections: for example, a tomato is connected to a class of vegetables, not garments.

RuWordNet, a Russian thesaurus, presents the concept of “vegetables” as a member of a synonymic set, a set of hyponyms and a set of hyperonyms (<https://ruWordnet.ru/ru/>).

The WordNet thesaurus is another tool widely used to measure semantic proximity of words. For example, by September 25, 2019, Google Scholar registered 58900 articles mentioning WordNet as a tool and word similarity as a research objective. Over 3150 works on the topic appeared in Google Scholar only in September 2019.

In [17], the authors offer the first systematic comparison of various measures of word proximity based on WordNet. In [18] the authors offer a broad survey of proximity metrics such as path length based measures, information content-based measures, feature-based measures, and hybrid measures. The innovative information-theoretic approach to measure the semantic similarity between concepts of WordNet is developed in [19]. In [20], it was proposed to assign weights to the edges of WordNet and determine the proximity of words based on “weighted WordNet”.

¹<https://wikidata.org/>

A comprehensive review of Russian thesauri is presented in [21]. It includes, in particular, information about RuThes and RuWordNet thesauri, i.e., Russian WordNet. In [22, 23], these thesauri are used to solve the problem of establishing semantic proximity of words.

To the best of our knowledge, research on the degree of conceptual coherence of a text, and hence its complexity implementing this approach is still scarce. In article [6], DBpedia was used as a knowledge base, and Newsela article corpus (<https://newsela.com/data>) as a corpus of texts, containing original and artificially simplified texts. Entities expressed in the text by nominal groups were mapped onto knowledge base concepts. All DBpedia concepts, together with the semantic relationships linking them to each other, form a graph. As a result, the text is displayed on a subgraph of the complete DBpedia graph. The authors of the article reviewed 13 parameters of the graph and calculated their values for original and simplified texts (total number of texts = 200). The research shows that all the parameters studied have a statistically significant relationship between metrics of the parameters of the graph and text complexity (at least in the case of significant differences in text complexity). Thus, this approach is viewed as reliable to assess texts complexity. In [24] the same authors propose a mechanism of distributing activation in a network of concepts which may be implemented to model the effect of priming. As priming is viewed as a mechanism accelerating text comprehension [25], it offers researchers another instrument to evaluate conceptual complexity of texts.

In this paper, we propose a novel approach to assess the conceptual complexity of texts. It differs from the approach proposed in [6] in many significant aspects: (1) implementing a WordNet-like thesaurus as a knowledge base, rather than a DBpedia knowledge base; (2) applying a set of structural features of the graph; (3) using natural texts of different conceptual complexity for testing the approach rather than artificially generated texts.

3. Materials and Methods

3.1. Datasets

For the experiment with English, we use two datasets, the Newsela corpus and the Simple English Wikipedia corpus. The Newsela corpus contains the data of 1130 news articles. Each article has 5 versions (1 original text and 4 simplified versions). Thus, this dataset can be used in a multiclass classification task. Another corpus is the Simple English Wikipedia that contains some of Wikipedia articles written primarily in Basic English. The data used in this study in a binary classification setting.

For experiment with Russian texts, we use Russian Readability Corpus (RRC). Russian Readability Corpus compiled for the current research comprises three sets of books, i.e. Social Science textbooks, History textbooks, Elementary school texts. Initially [4], RRC was compiled of two sets of textbooks on Social studies for secondary and high school for Russian students. It contained 45380 sentences from 14 textbooks: edited by Bogolyubov (BOG) and by Nikitin (NIK). Later, a dataset of 17 elementary school texts (1st – 4th grades) along with a dataset of 6 textbooks on History (10th – 11th grades) were added.

3.2. Linguistic resources and knowledge bases

In experiments with English texts, we applied two types of knowledge graphs: WordNet and Wikidata; while for Russian, we use RuThes-Lite Thesaurus. RuThes Thesaurus of the Russian language [26] is typically referred to as a linguistic knowledge base for natural language processing. The thesaurus provides a hierarchical network of concepts. Each concept has a name and is related to other concepts and to a set of language signs (words and phrases), the meanings of which correspond to the concept.

The conceptual relations in RuThes include the following:

- the class-subclass relation;
- the part-whole relation;
- the external ontological dependence, and others

RuThes contains 54 thousand concepts, 158 thousand unique text entries (75 thousand single words), 178 thousand concept-text entry relations, over 215 thousand conceptual relations. The first publicly available version of RuThes (RuThes-lite²). The process of generating RuThes-Lite from RuThes is described in [27]. For the goals of the present study we have transformed RuThes-Lite into a graph where vertices are concepts and edges are relationships (we use the class-subclass, the part-whole relations as well as association relation).

3.3. A method for generating a graph from a text

As it was mentioned above, for this research, we use knowledge graphs to estimate text complexity. The structure of a thesaurus is represented as a graph (G_0): with nodes derived from the concepts and edges derived from relations. While processing a text (or a fragment of a text) words from a text are matched with knowledge graph nodes concepts.

Performing a proper matching of thesaurus concepts with raw text entries is an important step that usually involves disambiguation processes. However, the problem of automatic disambiguation of a word sense in Russian has not been solved yet. Thus, when matching RuThes-Lite concepts with a text, we use a simple string, thus matching normalized words from the text and the text entries that correspond to the thesaurus concepts. We keep all the matching concepts in a temporary list. This process may produce a lot of false positives in the temporary list, i.e., concepts that were not used in the text fragment. On the next step, i.e., while building a subgraph, we filter out all isolated nodes. It is during this procedure the vast majority of false positives are excluded. In contrast, in experiments with English texts, we make use of well-developed disambiguation techniques that map words to WordNet's synsets.

The procedure of building a subgraph is straightforward: we use all matched concepts to produce a new graph G_S as a subgraph of G_0 . If two nodes are connected in G_0 , then they are connected in G_S too. In case of a false positive match, the false positive will remain in subgraph G_S if and only if it has a connection with another falsely matched concept in the given text. Consequently, having two interconnected false positives is still possible, but having more

²<http://www.labinform.ru/ruthes/index.htm>

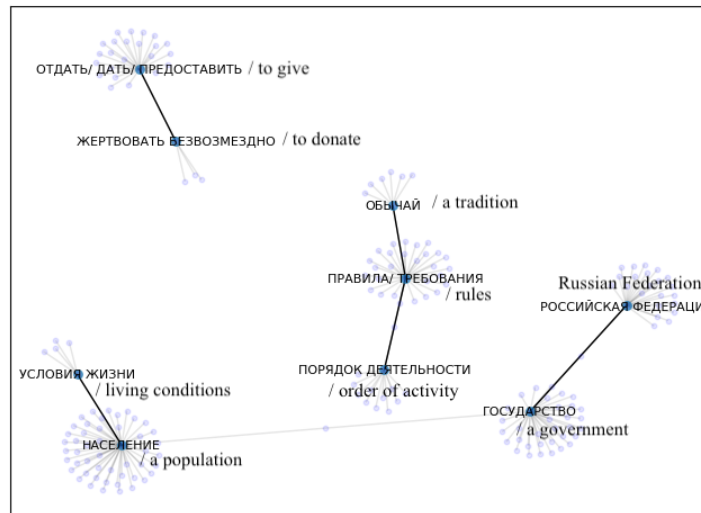


Figure 1: A subgraph G_S derived from the RuThes-Lite and a text fragment. The subgraph contains only nodes and (hypernymy-hyponymy) edges between them (such edges are dark). Nodes without labels are kept in the figure as an illustration of the related RuThes concepts that did not appear in the text fragment, but they are linked to nodes from the subgraph G_S

than three (interconnected false positives) is a much rarer event. Limitations of the described procedure are obvious as some of the isolated nodes may still contain valuable information for further analysis.

For example, a sample text segment from the 6-th grade textbook will be mapped to the corresponding subgraph of RuThes-Lite (Fig. 1); Figure 2 depicts a subgraph derived from a fragment from the Newsela corpus using WordNet.

3.4. Sampling and Graph-based features

RRC contains 37 documents and thus can hardly be viewed as a representative sample of the population of all school textbooks. However, for the purposes of text complexity studies, we could split each document in multiple non-overlapping parts. If each part (or a sample) is 'long enough' it can serve as a good representative of the whole document and at the same time will keep certain variability. As we have no assumptions on the idea of "big enough" in terms of the sample, we will denote the size of each sample as parameter S . This parameter is measured in tokens.

In experiments, both dependent variables (such as readability value) and independent variables measured for a given sample. The sample size (S) was set to different values: 200, 500, 1000 and 2000 tokens. During sampling, keeping the order of tokens and sentences is important; otherwise, the sampled texts will be less natural, even though they could carry the main features of the documents from the corpus. Thus, we sample S token sequences from each document³.

³The last sentence is not truncated, hence the size of a sample in experiments is at least S tokens and at most $(S+k)$ tokens, where k tokens are used to keep the last sentence in the sample

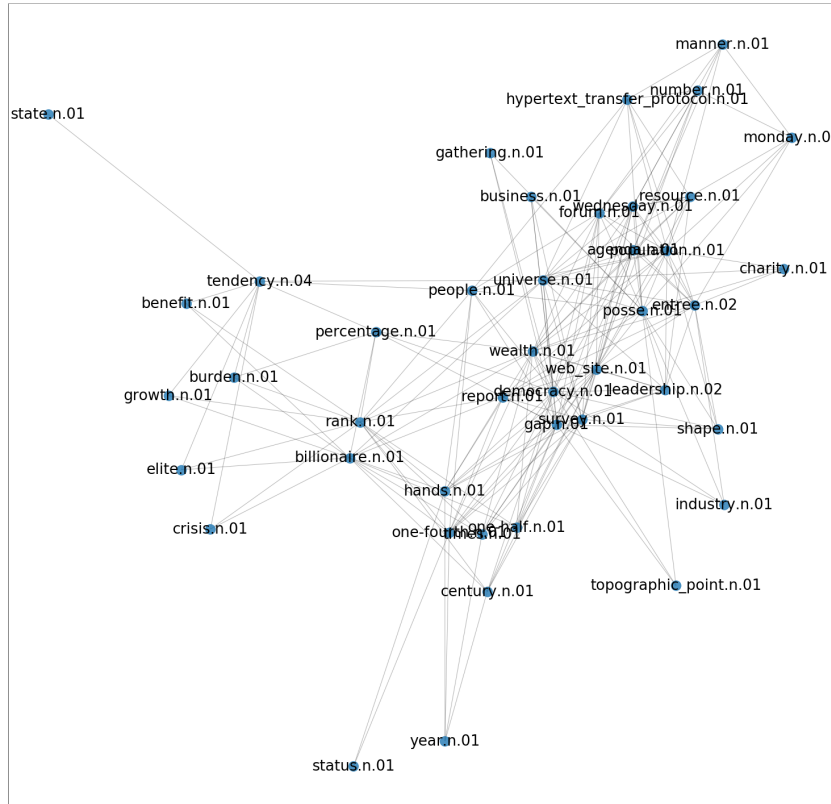


Figure 2: A subgraph G_S derived from the WordNet and a text fragment.

We calculate all features for readability analysis using the described sampling technique. Using the technique, we can estimate the mean and range of feature metrics with several samples taken from RRC. Sampling from RRC produces a set of text fragments (a text sample and a text fragment are used as synonyms). A subgraph G_S is generated from each sample.

Previously in [28] we conducted experiments with RRC data to investigate the correlation between thesaurus-based features and complexity. Below, we describe features that we use in experiments with text complexity and provide results from [28]. In our experiments, for each G_S we calculate the following features:

- number of RuThes concepts (C_o);
- number of components (NC);
- average component size (CS);
- maximum node degree (MND);
- total number of connected nodes (TCN);
- average shortest path (ASP).

Table 1

Pearson correlation between features based on the RuThes-Lite and the grade level

S	Co	NC	CS	MND	TCN	TCN / Co	ASP
200	0.45	0.86	0.32	0.53	0.84	0.76	0.27
500	0.62	0.86	0.68	0.45	0.88	0.79	-0.25
1000	0.56	0.38	0.73	0.40	0.85	0.69	-0.36
2000	0.88	0.40	0.81	0.68	0.86	0.70	-0.54

The correlation between features based on RuThes-Lite and the grade level presented in Table 1. In this table we highlight the values higher than 0.7 of the correlation coefficient. For each row of Table 1 we measured 50 random samples per a textbook from the RRC. In the next section, we present the results of experiments with English texts.

4. Experiments

The set of features that correlate with text complexity in previous experiments (in Russian) was also tested in English. For experiments in English texts we chose a text classification task; so the set of graph features were tested on a different language. We apply the technology described above to texts in English in order to test it.

Experiments for English texts were carried out on NEWSLA and Simple Wikipedia corpora. WordNet and Wikidata were used as knowledge bases. Disambiguation methods developed for the English language enable text annotation of high quality using WordNet concepts and methods for linking Wikidata entities to text (entity linking). Experiments have been carried out with classical machine learning methods (stochastic gradient descent, logistic regression, KNN, decision trees, and SVMs), as well as neural network graph models and graph embeddings.

4.1. Classification of texts by complexity based on the knowledge graph of WordNet

In experiments with classical machine learning methods, feature vectors were constructed with the help of specialized methods. In the beginning, each node from the WordNet hierarchy gets its own representation in the form of a vector (which is done by node2vec method [29]). The classifiers were trained on SimpleWikipedia data and tested on a binary classification problem. The dimension of embeddings for WordNet concepts was chosen equal to 128. The classification accuracy reached 93%. When two additional features were added (the average word length or average sentence length, the accuracy increased to 95%). At the same time, the model trained exclusively on the "classical" features showed the accuracy of no more than 85%. Similar (in terms of accuracy) results were obtained in [30]. In this group of experiments with neural network graph models, feature vectors were also built using node2vec. The resulting vectors were assigned to the vertices of the graph G , after which a graph convolution network (GCN, [31]) was applied to the graph. We also used GCNs to construct a convolution of graph G and extract features directly. The classifier was trained on data from the Simple Wikipedia corpus and

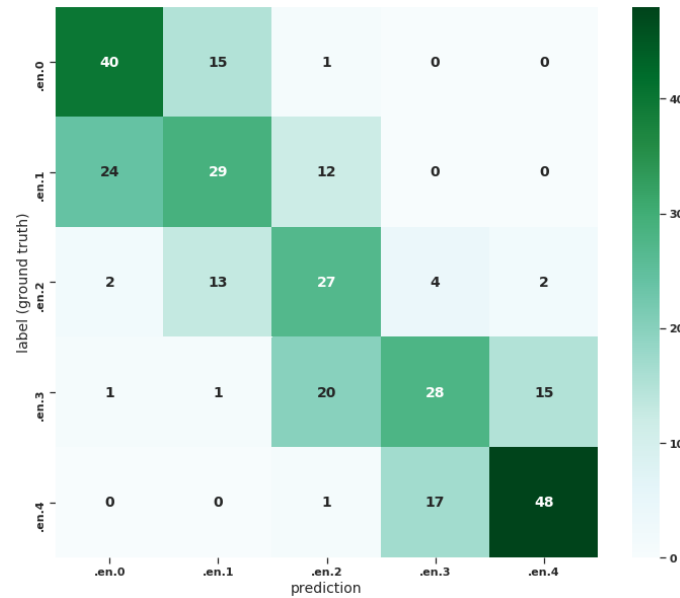


Figure 3: Confusion matrix of Decision tree classifier on graph-based features (the Newsela dataset).

tested on a binary classification. The dimension of embeddings for WordNet concepts equals 256. The classification accuracy on the test sample received was 60 – 65%. Parameters change did not result in improving the model quality.

4.2. Classification of texts by complexity based on the Wikidata knowledge graph

To extract Wikidata concepts from text, researchers in [32] implemented the BLINK model. Pretrained vectors of dimension 200 (Pytorch-BigGraph [33]) were used as vector representations for the nodes of the Wikidata graph. The classifier was trained on NEWSLA dataset. The maximum classification accuracy obtained was 62.5% Parameters change did not result in improving the model quality (Fig. 3).

General conclusions on the use of thesauri: classical machine learning models based on graph features demonstrated better results compared to graph convolution neural networks [31]. Models based on only "superficial" features (such as average word length or average sentence length, FOG, SMOG, and others derived from TextStat⁴ library) can be improved by adding vector representations trained on WordNet graph. However, the use of vectors built on Wikidata graph requires further research to analyse the sources of very moderate performance.

⁴<https://pypi.org/project/textstat/>

5. Conclusion

Text complexity is of utmost importance both for textbook authors and students looking for educational materials. Modern methods and approaches of artificial intelligence, including knowledge bases, allow assessing conceptual complexity of texts, thus providing educators and students with instruments they need. The combined application of methods of computational linguistics and artificial intelligence can be successfully used when determining text complexity and thus contribute to significant progress in understanding the notion of text complexity at a deeper conceptual level.

In the paper, we presented and evaluated graph-based complexity features. Such features can be extracted from text fragments using hierarchical structure of a thesaurus. A previous study was conducted on the material of the Russian texts and RuThes-Lite thesaurus only. We have evaluated correlation of the features with text complexity.

The present work deals with texts in English. It is expected that similar results can be achieved for other languages with the help of the same methods as on the conceptual level, languages reflect the real world using the same cognitive mechanisms. Switching to another language needs changing the corresponding linguistic ontology. Therefore, we applied a similar approach using WordNet.

Despite the results are promising, there are still many research questions still open in the area. The perspective of the study lies in the detailed analysis of using knowledge graphs such as Wikidata and comparing the effectiveness of results derived with WordNet. While the current study shows somewhat mediocre performance of graph convolution networks, an increased number of features and their automatic selection could be very fruitful to detect relevant features of text complexity. Our work in a number of aspects can be treated as a benchmark applicable in further studies of conceptual complexity of texts.

Acknowledgments

The study presented in paper has been supported by the Kazan Federal University Strategic Academic Leadership Program (Sections 1–3 of the paper), and by the Russian Science Foundation, grant 18-18-00436 (Sections 4–5 of the paper).

References

- [1] M. Solnyshkina, A. Kisel'nikov, Slozhnost' teksta: etapy izucheniya v otechestvennom prikladnom yazykoznanii [text complexity: study phases in Russian linguistics], Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya [Tomsk State University Journal of Philology] (2015).
- [2] V. Ivanov, M. Solnyshkina, V. Solovyev, Efficiency of text readability features in Russian academic texts, Komp'yuternaja Lingvistika i Intellektual'nye Tehnologii 17 (2018) 277–287.
- [3] R. Reynolds, Insights from Russian second language readability classification: complexity-dependent training requirements, and feature evaluation of multiple cate-

- gories, in: Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications, 2016, pp. 289–300.
- [4] V. Solovyev, V. Ivanov, M. Solnyshkina, Assessment of reading difficulty levels in Russian academic texts: Approaches and metrics, *Journal of Intelligent & Fuzzy Systems* 34 (2018) 3049–3058.
- [5] B. Biryukov, B. Tyukhtin, O ponyatii slozhnosti [about the concept of complexity], V kn.: *Logika i metodologiya nauki. Materialy IV Vsesoyuznogo simpoziuma.* (1967) 219–231.
- [6] S. Štajner, I. Hulpus, Automatic assessment of conceptual text complexity using knowledge graphs, in: Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics, 2018, pp. 318–330. URL: <http://aclweb.org/anthology/C18-1027>.
- [7] C. A. Denton, M. Enos, M. J. York, D. J. Francis, M. A. Barnes, P. A. Kulesz, J. M. Fletcher, S. Carter, Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text, *Reading Research Quarterly* 50 (2015) 393–416.
- [8] D. S. McNamara, A. Graesser, M. M. Louwerse, Sources of text difficulty: Across genres and grades, *Measuring up: Advances in how we assess reading ability* (2012) 89–116.
- [9] Y. A. Tomina, Ob’ektivnaya otsenka yazykovoy trudnosti tekstov (opisanie, povestvovanie, rassuzhdenie, dokazatel’stvo)[an objective assessment of language difficulties of texts (description, narration, reasoning, proof)], Abstract of Pedagogy Cand. Diss. Moscow (1985).
- [10] A. Laposhina, Relevant features selection for the automatic text complexity measurement for Russian as a foreign language, *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”* (2017) 1–7.
- [11] D. McNamara, A. Graesser, P. McCarthy, Z. Cai, *Automated evaluation of text and discourse with Coh-Metrix*, Cambridge University Press, 2014.
- [12] S. Crossley, D. McNamara, Text coherence and judgments of essay quality: Models of quality and coherence, in: Proceedings of the Annual Meeting of the Cognitive Science Society, volume 33, 2011.
- [13] S. Thornbury, Beyond the Sentence: Introducing Discourse Analysis, *ELT Journal* 60 (2006) 392–394. URL: <https://doi.org/10.1093/elt/ccl033>. doi:10.1093/elt/ccl033. arXiv:<http://oup.prod.sis.lan/eltj/article-pdf/60/4/392/1346812/ccl033.pdf>.
- [14] A. Collins, M. Quillian, Retrieval time from semantic memory, *Journal of Verbal Learning and Verbal Behavior* 8 (1969) 240–248.
- [15] M. Steyvers, J. B. Tenenbaum, The large-scale structure of semantic networks: statistical analyses and a model for semantic growth, *Arxiv preprint cond-mat/0110012* (2001).
- [16] C. Fellbaum (Ed.), *WordNet: an electronic lexical database*, MIT Press, 1998.
- [17] A. Budanitsky, G. Hirst, Evaluating wordnet-based measures of lexical semantic relatedness, *Computational Linguistics* 32 (2006) 13–47.
- [18] L. Meng, R. Huang, J. Gu, A review of semantic similarity measures in wordnet, *International Journal of Hybrid Information Technology* 6 (2013) 1–12.
- [19] T. Hong-Minh, D. Smith, Word similarity in wordnet, in: *Modeling, Simulation and Optimization of Complex Processes*, Springer, 2008, pp. 293–302.
- [20] M. G. Ahsae, M. Naghibzadeh, S. E. Y. Naeni, Semantic similarity assessment of words

using weighted wordnet, *International Journal of Machine Learning and Cybernetics* 5 (2014) 479–490.

- [21] N. S. Lagutina, K. V. Lagutina, A. S. Adrianov, I. V. Paramonov, Russian language thesauri: automated construction and application for natural language processing tasks, *Modelirovanie i Analiz Informatsonnykh Sistem* 25 (2018) 435–458.
- [22] N. Loukachevitch, A. Alekseev, Summarizing news clusters on the basis of thematic chains, in: *Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, 2014, pp. 1600–1607.
- [23] D. Ustalov, Concept discovery from synonymy graphs, *Vychislitel'nye tekhnologii [Computational Technologies]* 22 (2017) 99–112.
- [24] I. Hulpus, S. Štajner, H. Stuckenschmidt, A spreading activation framework for tracking conceptual complexity of texts, in: *Proceedings of the 57th Conference of the Association for Computational Linguistics*, 2019, pp. 3878–3887.
- [25] T. Gulan, P. Valerjev, Semantic and related types of priming as a context in word recognition, *Review of psychology* 17 (2010) 53–58.
- [26] N. Loukachevitch, B. Dobrov, RuThes linguistic ontology vs. Russian wordnets, in: *Proceedings of the Seventh Global Wordnet Conference*, 2014, pp. 154–162.
- [27] N. Loukachevitch, B. Dobrov, I. Chetviorkin, Ruthes-lite, a publicly available version of thesaurus of Russian language ruthes, in: *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”*, volume 2014, 2014.
- [28] V. Solovyev, V. Ivanov, M. Solnyshkina, Thesaurus-based methods for assessment of text complexity in russian, in: *Mexican International Conference on Artificial Intelligence, Springer*, 2020, pp. 152–166.
- [29] A. Grover, J. Leskovec, node2vec: Scalable feature learning for networks, in: *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 2016, pp. 855–864.
- [30] Z. Jiang, Q. Gu, Y. Yin, D. Chen, Enriching word embeddings with domain knowledge for readability assessment, in: *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 366–378.
- [31] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, *arXiv preprint arXiv:1609.02907* (2016).
- [32] L. Wu, F. Petroni, M. Josifoski, S. Riedel, L. Zettlemoyer, Scalable zero-shot entity linking with dense entity retrieval, *arXiv preprint arXiv:1911.03814* (2019).
- [33] A. Lerer, L. Wu, J. Shen, T. Lacroix, L. Wehrstedt, A. Bose, A. Peysakhovich, Pytorch-biggraph: A large-scale graph embedding system, *arXiv preprint arXiv:1903.12287* (2019).