Text Simplification for Lithuanian

Justina Mandravickaitė[0000−0001−9426−6165], Eglė Rimkienė[0000−0002−9352−9709], Danguolė Kotryna Kapkan[0009−0007−1673−9626], and Danguolė Kalinauskaitė[0000−0003−0641−2324]

Vytautas Magnus University, Kaunas, Lithuania

Abstract. Text simplification involves reducing complexity while retaining important information. This is important for improving accessibility for a wide range of readers, including those with cognitive disorders, non-native speakers, and children. We report experiments on text simplification for Lithuanian with a focus on simplifying texts, written in an administrative style, which is not easy to understand for the general public. We chose mT5 as a foundational model and fine-tuned it in 3 different ways. We evaluated models' outputs with 'unseen' data and automatic evaluation metrics accompanied by qualitative analysis. Though automatic evaluation scores were higher for the model fine-tuned with translated data, qualitative analysis revealed that the model that was fine-tuned with specifically for this task prepared Lithuanian data performed simplification task significantly better.

Keywords: Text simplification · Lithuanian · Transformers · Fine-tuning.

1 Introduction

Text simplification is the task of simplifying the vocabulary and syntactic complexity of a text while preserving the essential information of the original text. Text simplification is relevant for improving the accessibility of information for people with cognitive disorders, as well as for non-native speakers and children [30].

In this paper, we report text simplification experiments for Lithuanian. We focus on the simplification of institutional texts, specifically written in the administrative (clerical) style. The examples of communication with the public by public authorities often use quasi-legal language, which can be ineffective in conveying information to non-specialists. This makes the texts difficult to understand for anyone who is not an expert in the subject matter. While public communication texts on the websites of various public administration institutions are intended to disseminate information of public interest, such as social benefits, public utilities, migration, copyright, and other administrative issues, there is quite often a contradiction between their declared purpose and actual functionality. The texts on public institutions' websites may be too complex, making it unclear whether they are intended for the general and can meet their needs for information on important issues, problems, and procedures. Text simplification has the potential to address this problem as it transforms administrative language into a less complex one in terms of vocabulary, sentence structure, and other aspects, while at the same time retaining the essential information from the original content.

We chose mT5 as a base model and fined-tuned it in 3 ways, developing 3 models for simplifying Lithuanian texts. The rest of the paper is structured as follows: Section 2 briefly introduces related work, Section 3 describes data used for training and evaluation, Section 4 – methods used in our experiments, Section 5 – experimental setup, Section 6 presents results. Finally, Section 7 ends the paper with conclusions.

2 Related Work

The popular methods for text simplification involve using pre-trained language models to control the desired level of simplicity of the output. As transformers' attention architecture looks at the entire input sequence and selectively extracts essential information [34], this feature has been found advantageous for text simplification (e.g. [38, 22]), among other NLP tasks. BERT model has been used for lexical text simplification (e.g., [25]), text simplification using monolingual machine translation [3], or hybrid text simplification approach (e.g., [18]), among other studies. The T5 model has been used for controllable text simplification (e.g., [29]) as well as for text simplification in a situation with limited resources (e.g., [20]), among others.

Recent studies have demonstrated that these models can simplify text via different techniques, such as specifying the desired reading grade level or directly indicating necessary simplification operations [1]. Some of the newest models for text simplification include SIMSUM for automated document-level text simplification [7], also, SimpleBART [31], which reports a pre-training strategy for text simplification, and KGSimple, an unsupervised approach that uses knowledge graphs to generate compressed text [9]. In addition to general text simplification, domain-specific text simplification models are emerging, e.g., for simplifying medical texts [6].

What makes text simplification a complex and non-trivial task, is the lack of high-quality data sources and the need for further exploration of the low-resource scenarios [32]. Additionally, sometimes domain-specific text simplification may result in lower quality generated text as on, e.g., medical text simplification [16, 13]. Finally, there are challenges related to cultural and commonsense knowledge in text simplification which requires further research in this field [10].

3 Data

3.1 Training Data

Lithuanian Dataset for Text Simplification. The final dataset comprises 2,142 entries with two columns, where the first column contains original sentences or text fragments, equivalent to sentences, while the second column contains manually simplified versions of the corresponding original text content. All data were simplified by four experts according to guidelines drawn up based on the literature on plain language, i.e. simplified version of language, intended for non-specialists (general public) [2].

The data sources for this study were texts from various Lithuanian governmental and non-governmental public institution websites that provide information on services such as social benefits, migration, utilities, copyright, and other administrative issues. The simplification process involved dividing the texts into sentences or sentence-equivalent text fragments and transforming them manually following a set of predefined simplification rules.

The lexical and syntactic rules used for simplification were primarily based on cross-linguistic *Plain Language* principles [15, 19]. In some instances, we also considered *Plain Language* principles or text simplification syntactic rules that are specific to languages that are structurally similar to Lithuanian [12, 8]. Also, we defined certain rules, such as the treatment of participles, specifically for the Lithuanian language. For lexical simplification, word frequency, as indicated in the 'Frequency Dictionary of the Written Lithuanian Language' [33], was taken into account when in doubt, as the simpler words are also more frequent ones.

Translated Datasets for Text Simplification. Besides parallel corpus for Lithuanian text simplification, we used additional English datasets developed for the same task which we translated into Lithuanian via machine translation. We chose the following datasets for our experiments:

1. **MASSAlign_philo**[1]: this dataset provides texts of philosophical discourse [24]. The data aligned the original texts, obtained from *Project Gutenberg*[2] with their simplified versions that were edited and retrieved from *Early-ModernTexts*[3]. The alignment was performed via MASSAlign technique for aligning comparable documents [23]. The dataset consists of 8453 aligned sentences.
2. **SimPA**[4]: it is a corpus designed for sentence-level simplification in the domain of public administration. It consists of 1,100 original sentences, each of which was manually simplified in a two-step process - lexically and then syntactically [27]. The dataset contains 3,300 lexically simplified sentences and 1,100 syntactically simplified sentences.
3. **Human Simplification with Sentence Fusion Data Set (HSSF)**[5]: The dataset consists of three separate files, each intended for different research and development purposes. We utilized the *training* subset of this data, which includes 119 original sentences. Each sentence is accompanied by four

[1] https://github.com/stefanpaun/massalign
[2] https://www.gutenberg.org/
[3] https://www.earlymoderntexts.com/
[4] https://github.com/SIMPATICOProject/simpa
[5] https://cs.pomona.edu/ dkauchak/simplification/

human-generated simplifications and ten fusion sentences [28]. Simplifications made by humans were obtained through Amazon Mechanical Turk[6]. The simplicity and adequacy scores, which were derived from the dataset, represent the average evaluations of three individual workers. [28].

3.2 Evaluation Data

For the evaluation of our models, we used a test sample of 102 sentences from texts published on various websites of Lithuanian governmental and non-governmental public institutions. These sentences were not included in the training data, therefore, they were 'unseen' by the models. All the models were evaluated using this dataset. Evaluation includes scores of automatic evaluation metrics and qualitative assessment of models' outputs (see Section 4).

4 Methods

In our study, we focus on the task of text simplification for the Lithuanian language. Our approach leverages advanced natural language processing (NLP) techniques, specifically the adaptation and fine-tuning of the mT5 (Multilingual Text-to-Text Transfer Transformer) model, to address the unique linguistic features of Lithuanian. The results were evaluated via automatic evaluation metrics as well as qualitative analysis.

4.1 Models

The foundation of our methodology is based on the T5 model, which stands for "Text-to-Text Transfer Transformer." Developed by Google, T5 adopts a unified text-to-text framework, where every language processing task is re-framed as a text generation problem. Key principles of the T5 model include [36]:

1. **Unified Text-to-Text Framework**: T5 treats all NLP tasks as a text generation problem, where the input and output are always text strings. This approach simplifies the architecture and allows for flexibility in handling NLP tasks.
2. **Pre-training on a Diverse Corpus**: T5 is pre-trained on a large, diverse corpus, C4 (Colossal Clean Crawled Corpus) [11], which provides a broad understanding of language and context.
3. **Encoder-Decoder Architecture**: The model employs an encoder-decoder architecture, similar to the original Transformer model. The encoder processes the input text and creates a contextual representation, which the decoder then uses to generate the output text.
4. **Fine-Tuning for Specific Tasks**: While T5 is pre-trained on a general corpus, it can be fine-tuned on a specific task or language to enhance its performance.

[6] https://www.mturk.com/

For our specific task of Lithuanian text simplification, we utilized the mT5 model, a multilingual variant of the original T5 [35]. This model is pre-trained on a multilingual dataset covering 101 languages, making it particularly suitable for languages with fewer resources, such as Lithuanian.

4.2 Evaluation

Metrics

- **BLEU (Bilingual Evaluation Understudy) Score**: It measures the number of matching n-grams between the output and reference sentences. Initially developed for evaluating machine translation, it is also used for evaluating text simplification and other NLP tasks as it is easy to calculate and interpret [26].
- **ROUGE (Recall-Oriented Understudy for Gisting Evaluation) Score**: It measures the overlap of n-grams between the simplified text and reference text in different ways depending on selected units for comparison, e.g., unigrams, bigrams, etc. [17]. We chose 3 variants of ROUGE: unigram overlap (ROUGE-1), bigram overlap (ROUGE-2) and Longest Common Subsequence overlap (ROUGE-L).
- **ROLD (Replace-only Levenshtein distance)**: The Levenshtein distance is a metric used to measure the dissimilarity between two string sequences. The text is characterized by the minimum number of operations required to transform one string into another, including insertions, deletions, or substitutions[4]. We selected this metric to assess the performance of models on a symbol level.
- **BERTscore**: It identifies similar words in candidate and reference sentences using cosine similarity with pre-trained contextual embeddings from BERT. It correlates well with human evaluation and its range is between -1 and 1 [37].

Qualitative Analysis To evaluate the simplification output qualitatively, three criteria are commonly used: grammaticality, meaning preservation, and simplicity [21, 5]. Grammaticality refers to whether the simplified text remains grammatical and understandable. Meaning preservation evaluates whether the semantics or adequacy is preserved after simplification. Simplicity assesses whether the simplified text is simpler than the original text. These criteria are evaluated without the need for reference data [14].

The expert has been asked to assess sentences, simplified by our 3 models according to these 3 criteria on a scale from 1 to 5. As all 3 evaluation criteria are not equal (they go in this order: simplicity – meaning retention – grammaticality), we also asked to apply 2 other rules in the evaluation:

- The most important criterion is *simplicity*, so if according to this criterion generated sentence gets 1, meaning retention and grammaticality are irrelevant (gets the score of 1).

– If for *simplicity* a sentence scores higher than 1, but *meaning retention* scores 1, then the grammaticality is scored 1 (otherwise we would get a grammatically correct sentence which is semantically unrelated to the original one).

Without such a hierarchy of criteria, we would have a paradoxical situation where models would be rewarded for doing nothing to the text, just copying it, and penalizing those that attempt to simplify, although with errors.

5 Experimental Setup

Structured Training Process

The mT5 model's ability to simplify Lithuanian text was enhanced through a structured training process. This process involved different datasets and was divided into several key experiments as follows:

1. **Fine-Tuning with Parallel Corpus for Lithuanian Text Simplification:** The first experiment used a dataset consisting of 2,142 Lithuanian sentences in both original and simplified forms. This was essential for teaching the model to understand and generate simplified Lithuanian text, ensuring the preservation of the original text's meaning.
2. **Fine-Tuning with Translated Data and Lithuanian Parallel Corpus:** The model's learning was further enhanced by adding 10,445 sentences translated from English to Lithuanian, both in original and simplified forms. This additional dataset was combined with the initial one to provide a diverse range of text, with the expectation of improving the model's proficiency in Lithuanian text simplification tasks.
3. **Fine-Tuning with Translated Dataset Only:** The final experiment focused solely on the translated dataset to assess the model's performance on content translated into Lithuanian, which poses different linguistic challenges compared to native texts.

Each experiment was designed to incrementally improve the model's proficiency in simplifying Lithuanian text, considering various linguistic styles and contexts. The results of these experiments are presented in Section 6.

Training Hyperparameters

The training of our model was governed by several key hyperparameters, each chosen to optimize the learning process and model performance:

– **Learning Rate (0.0001):** A fundamental parameter, the learning rate dictates the speed at which our model updates its knowledge. The chosen value ensures a balance between fast convergence and avoiding overshooting the optimal solution.

- **Batch Size (4 for both training and evaluation):** Batch size determines the number of data points the model processes at one time. A size of 4 strikes a balance between computational efficiency and the ability to capture data variability.
- **Optimizer (Adam):** The Adam optimizer was selected for its effectiveness in handling sparse gradients and adaptive learning rates, making it ideal for our model's training.
- **Epochs (8):** The number of epochs defines how many times the entire dataset is passed through the model. Eight epochs were chosen to ensure sufficient learning without causing over-fitting.

These hyperparameters were chosen via experimentation to ensure the model learns efficiently and effectively, at the same time balancing the need for speed, accuracy, and generalization.

6 Results

6.1 Model Training Evaluation

While training the mT5 model on Lithuanian text simplification, we monitored several key performance indicators to evaluate the model's learning progression and efficiency (see Figure 1). The loss graph displayed a consistent decline, indicating successful learning and adaptation to the simplification task. Concurrently, ROUGE metrics (*eval/rouge1, eval/rouge2, and eval/rougeL*) were utilized to assess the quality of text simplification, with values reaching satisfactory levels, demonstrating the model's capability to generate high-quality, simplified text. Additionally, the *eval/gen_len* and *eval/runtime* graphs were examined to ensure the model produced simplified texts of reasonable length without excessive verbosity and that the training process remained stable over time. These metrics collectively provided a comprehensive understanding of the model's performance and established a robust experimental framework for the task at hand.

6.2 Model Output Evaluation

Comparison of Evaluation Metrics' Scores

For this evaluation, original sentences, which the models simplified, were not present in the training corpus, i.e., they were previously "unseen" by the models. Automatic evaluation scores are summarised in Table 1. Our mT5 model fine-tuned with translated data (*mT5_transl.*) had the best scores for all the metrics. This shows that generated output presents a high n-gram match between generated and reference sentences. Also, it reveals that the reference sentences and this model's generated sentences are the most similar on a symbolic level (ROLD = 193.77) as well as in terms of vector representations (BERTscore = 0.95).

Table 1: Automatic evaluation scores

Metrics	Model		
	mT5_lt	**mT5_transl.**	**mT5_lt+transl.**
BLEU ↑	0.23	**0.69**	0.66
ROUGE-1 (F_1) ↑	0.63	**0.91**	**0.91**
ROUGE-2 (F_1) ↑	0.55	**0.90**	0.89
ROUGE-L (F_1) ↑	0.61	**0.91**	**0.91**
ROLD ↓	345.22	**193.77**	203.30
BERTscore (F_1) ↑	0.80	**0.95**	**0.95**

The next best model is the one fine-tuned with the Lithuanian dataset as well as translated data (*mT5_lt+transl.*). This model shares ROUGE-1, ROUGE_L and BERTscore scores with *mT5_transl.* model. Finally, the model, fine-tuned with Lithuanian corpus only (*mT5_lt*) got the worst scores. That means that generated output presents a lower n-gram match between generated and reference sentences. Also, it shows that generated and reference sentences are rather different on the symbolic level (higher ROLD score) and in terms of vector representations (lower BERTscore score). To get a more comprehensive assessment, automatic evaluation via metrics was complemented with qualitative expert analysis, presented in the following subsection.

6.3 Qualitative Evaluation

As automatic evaluation does not cover all text simplification aspects, automatic evaluation has been accompanied by a qualitative evaluation by the expert, who assessed simplified sentences produced by all 3 models. The generated sentences were assessed by their simplicity, meaning retention and grammaticality. The results are summarised in Table 2.

Table 2: Qualitative Evaluation

	Simplicity	**Meaning retention**	**Grammaticality**
mT5_lt	**2,5686**	**2,2549**	**2,7941**
mT5_transl.	1,1471	1,1961	1,2549
mT5_lt+transl.	1,441	1,5294	1,7255

Upon qualitative assessment, we can see that *mT5_lt* model got the highest average scores for all 3 evaluation criteria. This contrasts with automatic evaluation scores, where this model got lower scores than the other 2 models. Automatic evaluation metrics we used measured a match between candidate and reference sentence on different levels – n-grams (BLEU and ROUGE), symbols (ROLD) and embeddings (BERTscore). However, qualitative analysis revealed that a significant proportion of sentences needed to be shortened. Instead of breaking down sentences by taking a relative clause, extracting its arguments and writing them as independent sentences, the mT5 tended to simply delete parts of sentences, losing bits of essential information and getting lower meaning retention scores. Automatic evaluation metrics failed to capture these and similar aspects of text simplification.

Also, the training data for *mT5_lt* was manually prepared parallel Lithuanian corpus for text simplification, where data was from the same domain as evaluation data. Not all translated data were from the administrative domain, e.g., HSSF was trained on news data, while MASSAlign_philo – philosophical texts. Also, translated and originally Lithuanian texts varied, at least to some degree, in cultural and societal contexts. Furthermore, English and Lithuanian differ significantly in terms of morphosyntactic aspects, which may have influenced the performance of our models and contributed to differences in the quality of text simplification. Finally, although the output of the model, fine-tuned on a combined Lithuanian dataset and translated data was evaluated better than the one, fine-tuned on the translated data only, data imbalance in training may have contributed to the results that need improvement.

7 Conclusions

We report experiments on text simplification for Lithuanian texts, written in an administrative style, which is not easy to understand for the general public. We chose mT5 as a foundational model and fine-tuned it in 3 different ways – 1) with Lithuanian parallel corpus for text simplification, 2) with translated data developed for text simplification task, and 3) with the data, combining the previous two. We evaluated models' outputs with "unseen" data and automatic evaluation metrics (BLEU, ROUGE, Levenshtein distance and BERTscore) accompanied by qualitative analysis (output assessment for simplicity, meaning retention and grammaticality). Though automatic evaluation scores were higher for the model fine-tuned with translated data, qualitative analysis revealed that the model that was fine-tuned with specifically for this task prepared Lithuanian data performed simplification task significantly better, scoring higher for all 3 evaluation criteria. Finally, the combination of Lithuanian and translated data did little to contribute to the quality of the output, possibly due to data imbalance during training.

Our future plans include model improvement and data augmentation to increase model performance and generalizability. Also, we plan a more comprehensive analysis of the model decision-making process to take into account such aspects as checking for factuality or model bias.

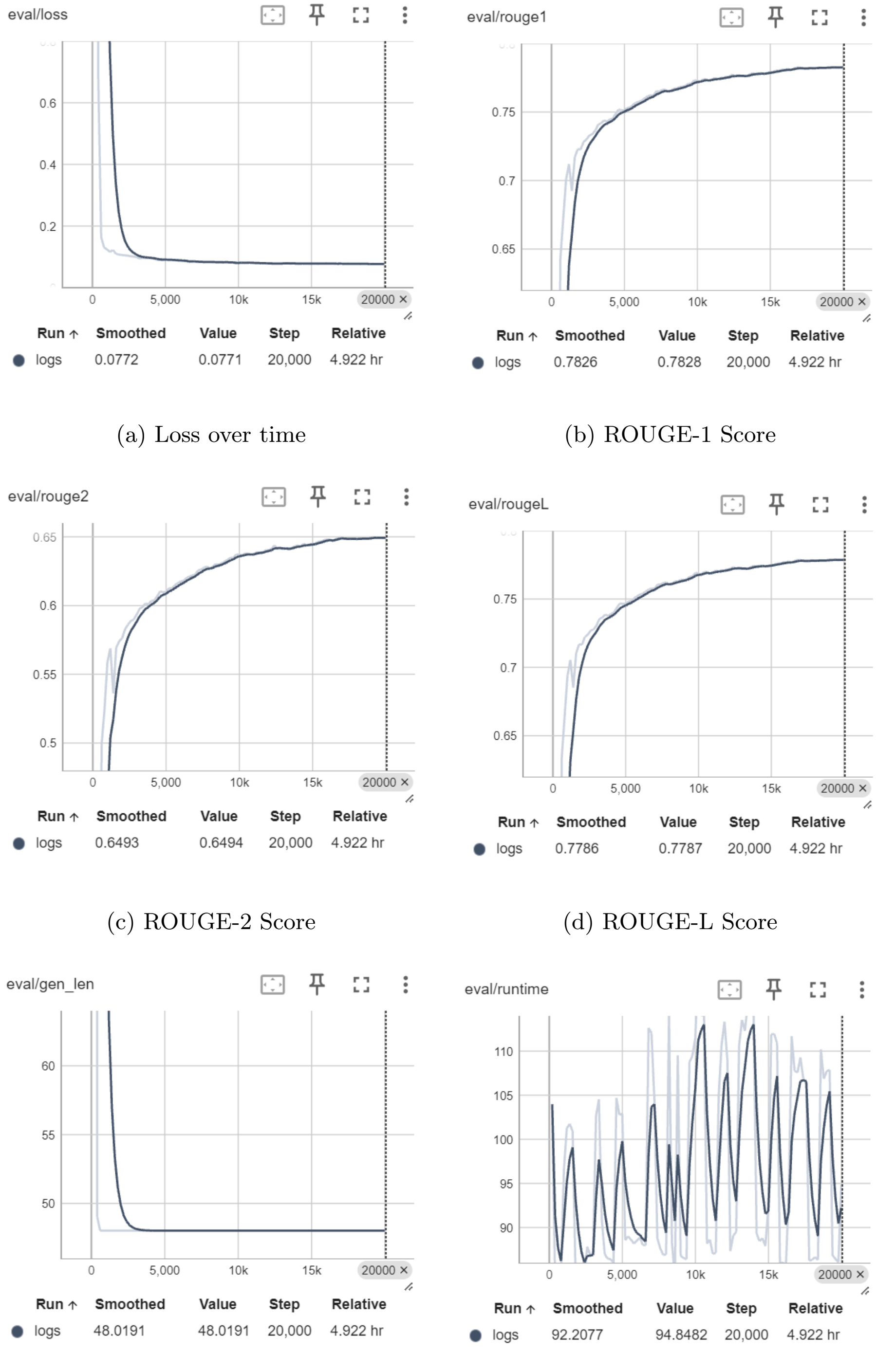
Acknowledgments. This research has received funding from the Research Council of Lithuania Lithuanian Studies Programme 2016–2024 under the project Automatic Adaptation of Administrative Texts in Lithuanian to the Needs of Non-Specialists, grant agreement No. S-LIP-22-77.

References

1. Agrawal, S., Carpuat, M.: Controlling pre-trained language models for grade-specific text simplification. In: Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. pp. 12807–12819 (2023)
2. Alarcon, R., Moreno, L., Martínez, P.: Lexical simplification system to improve web accessibility. IEEE Access **9**, 58755–58767 (2021)
3. Alissa, S., Wald, M.: Text simplification using transformer and bert. Computers, Materials & Continua **75**(2), 3479–3495 (2023)
4. Alva-Manchego, F., Martin, L., Bordes, A., Scarton, C., Sagot, B., Specia, L.: Asset: A dataset for tuning and evaluation of sentence simplification models with multiple rewriting transformations. In: ACL 2020-58th Annual Meeting of the Association for Computational Linguistics (2020)
5. Alva-Manchego, F., Scarton, C., Specia, L.: Data-driven sentence simplification: Survey and benchmark. Computational Linguistics **46**(1), 135–187 (2020)
6. Basu, C., Vasu, R., Yasunaga, M., Yang, Q.: Med-easi: finely annotated dataset and models for controllable simplification of medical texts. In: Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence. AAAI'23/IAAI'23/EAAI'23, AAAI Press (2023). https://doi.org/10.1609/aaai.v37i12.26649, https://doi.org/10.1609/aaai.v37i12.26649
7. Blinova, S., Zhou, X., Jaggi, M., Eickhoff, C., Bahrainian, S.A.: Simsum: Document-level text simplification via simultaneous summarization. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 9927–9944 (2023)
8. Brunato, D., Dell'Orletta, F., Venturi, G., Montemagni, S.: Design and annotation of the first Italian corpus for text simplification. In: Proceedings of the 9th Linguistic Annotation Workshop. pp. 31–41. Association for Computational Linguistics, Denver, Colorado, USA (Jun 2015). https://doi.org/10.3115/v1/W15-1604, https://aclanthology.org/W15-1604
9. Colas, A., Ma, H., He, X., Bai, Y., Wang, D.Z.: Can knowledge graphs simplify text? In: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. pp. 379–389 (2023)
10. Corti, L., Yang, J.: Artist: Artificial intelligence for simplified text. arXiv preprint arXiv:2308.13458 (2023)
11. Dodge, J., Sap, M., Marasović, A., Agnew, W., Ilharco, G., Groeneveld, D., Mitchell, M., Gardner, M.: Documenting large webtext corpora: A case study on the colossal clean crawled corpus. arXiv preprint arXiv:2104.08758 (2021)
12. Łukasz Dębowski, Bartosz Broda, B.N.E.C.: Jasnopis–a program to compute readability of texts in polish based on psycholinguistic research. Natural Language Processing and Cognitive Science (2015)
13. Flores, L.J.Y., Huang, H., Shi, K., Chheang, S., Cohan, A.: Medical text simplification: Optimizing for readability with unlikelihood training and reranked beam search decoding. In: Findings of the Association for Computational Linguistics: EMNLP 2023. pp. 4859–4873 (2023)
14. Grabar, N., Saggion, H.: Evaluation of automatic text simplification: Where are we now, where should we go from here. In: Traitement Automatique des Langues Naturelles. pp. 453–463. ATALA (2022)

15. Harris, Lynda, S.K..C.M.: Setting plain language standards. Clarity (2010)
16. Joseph, S., Kazanas, K., Reina, K., Ramanathan, V., Xu, W., Wallace, B., Li, J.J.: Multilingual simplification of medical texts. In: Bouamor, H., Pino, J., Bali, K. (eds.) Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. pp. 16662–16692. Association for Computational Linguistics, Singapore (Dec 2023). https://doi.org/10.18653/v1/2023.emnlp-main.1037, https://aclanthology.org/2023.emnlp-main.1037
17. Lin, C.Y.: Rouge: A package for automatic evaluation of summaries. In: Text summarization branches out. pp. 74–81 (2004)
18. Maddela, M., Alva-Manchego, F., Xu, W.: Controllable text simplification with explicit paraphrasing. arXiv preprint arXiv:2010.11004 (2020)
19. Martinho, M.: International standard for clarity - we bet this works for all languages. Clarity Journal (2018)
20. Monteiro, J., Aguiar, M., Araújo, S.: Using a pre-trained simplet5 model for text simplification in a limited corpus. Proceedings of the Working Notes of CLEF (2022)
21. Nisioi, S., Štajner, S., Ponzetto, S.P., Dinu, L.P.: Exploring neural text simplification models. In: Proceedings of the 55th annual meeting of the association for computational linguistics (volume 2: Short papers). pp. 85–91 (2017)
22. Omelianchuk, K., Raheja, V., Skurzhanskyi, O.: Text simplification by tagging. arXiv preprint arXiv:2103.05070 (2021)
23. Paetzold, G., Alva-Manchego, F., Specia, L.: Massalign: Alignment and annotation of comparable documents. In: Proceedings of the IJCNLP 2017, System Demonstrations. pp. 1–4 (2017)
24. Paun, S.: Parallel text alignment and monolingual parallel corpus creation from philosophical texts for text simplification. In: Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: student research workshop. pp. 40–46 (2021)
25. Qiang, J., Li, Y., Zhu, Y., Yuan, Y., Wu, X.: Lexical simplification with pretrained encoders. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34, pp. 8649–8656 (2020)
26. Reiter, E.: A structured review of the validity of BLEU. Computational Linguistics **44**(3), 393–401 (Sep 2018). https://doi.org/10.1162/coli_a_00322, https://aclanthology.org/J18-3002
27. Scarton, C., Paetzold, G., Specia, L.: Simpa: A sentence-level simplification corpus for the public administration domain. In: Proceedings of the LREC 2018 (2018)
28. Schwarzer, M., Tanprasert, T., Kauchak, D.: Improving human text simplification with sentence fusion. In: Proceedings of the 15th Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15). pp. 106–114 (2021)
29. Sheang, K.C., Saggion, H.: Controllable sentence simplification with a unified text-to-text transfer transformer. In: Proceedings of the 14th International Conference on Natural Language Generation (INLG); 2021 Sep 20-24; Aberdeen, Scotland, UK. Aberdeen: Association for Computational Linguistics; 2021. ACL (Association for Computational Linguistics) (2021)
30. Štajner, S.: Automatic text simplification for social good: Progress and challenges. ACL-IJCNLP 2021 pp. 2637–2652 (2021)
31. Sun, R., Xu, W., Wan, X.: Teaching the pre-trained model to generate simple texts for text simplification. In: Rogers, A., Boyd-Graber, J., Okazaki, N. (eds.) Findings of the Association for Computational Linguistics: ACL 2023. pp. 9345–9355. Association for Computational Linguistics,

Toronto, Canada (Jul 2023). https://doi.org/10.18653/v1/2023.findings-acl.595, https://aclanthology.org/2023.findings-acl.595
32. Sun, R., Yang, Z., Wan, X.: Exploiting summarization data to help text simplification. In: Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics. pp. 39–51 (2023)
33. Utka, A.: Dažninis rašytinės lietuvių kalbos žodynas. VDU leidykla (2009)
34. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
35. Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., Barua, A., Raffel, C.: mt5: A massively multilingual pre-trained text-to-text transformer. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 483–498 (2021)
36. Zhang, J., Zhao, H., Boyd-Graber, J.: Contextualized rewriting for text simplification. In: Transactions of the Association for Computational Linguistics. vol. 9, pp. 1525–1540 (2021). https://doi.org/10.1162/tacl_a_00369, https://doi.org/10.1162/tacl$_{a0}$0369
37. Zhang, T., Kishore, V., Wu, F., Weinberger, K.Q., Artzi, Y.: Bertscore: Evaluating text generation with bert. In: International Conference on Learning Representations (2019)
38. Zhao, S., Meng, R., He, D., Andi, S., Bambang, P.: Integrating transformer and paraphrase rules for sentence simplification. arXiv preprint arXiv:1810.11193 (2018)



(a) Loss over time

(b) ROUGE-1 Score

(c) ROUGE-2 Score

(d) ROUGE-L Score

(e) Generated Length

(f) Runtime

Fig. 1: Model performance metrics