AI-simplification of Mark Twain's *The Adventures of Tom Sawyer*: Assessment and Considerations

Adela Chindris[1], Madalina Chitez[2]

West University of Timisoara, Romania[1, 2]

Abstract

Text simplification, a linguistic concept dedicated to enhancing the understanding of written texts, is very important within educational contexts. The improvement of text understanding can be achieved through various strategies, such as grammatical, lexical or topical simplification. Presently, digital platforms, like ChatGPT, offer automatic text simplification, yet the efficacy of these simplifications remains untested. This paper aims to investigate the effectiveness of AI-generated text simplifications of some excerpts taken from Mark Twain's "The Adventures of Tom Sawyer" across multiple criteria, including readability scores and vocabulary complexity. The study uses widely accepted readability scores, such as Flesch-Kincaid and Coleman-Liau, and the online tool Coh-Metrix to assess the impact of simplification on text comprehension. Simplifications will be analyzed from the angles of syntactic complexity and lexical complexity. The study also includes an applied section featuring a questionnaire addressed to primary and middle schoolers, with the intent to assess the effectiveness of the simplifications.

1. Introduction

Written texts, whether they are novels, scientific articles or cordial letters, have properties that make them understandable to their readers. In order to decide whether a text is accessible to a particular group of readers, there are several readability metrics used to measure sentence and word lengths. Two of the most used readability scales are Flesch-Kincaid scale and Coleman-Liau scale and they do not evaluate how coherent a text it, but rather to whom it is suitable for, to which age gap it is intended (Siddharthan 2004, p. 18) [1]. Many educators, scientists and writers have considered adapting their articles, texts or handbooks to their audience. However, in some cases adjusting the readability of a text is more difficult and this is how text simplification has emerged. Text simplification (TS) is the process that aims to make a written text more accessible to its readers while preserving the meaning. There are many methods used to simplify text: "dis-embedding of relative clauses, the separation of subordinate clauses and coordinated verb phrases, the conversion from passive to active voice and the replacement of difficult words with easier synonyms" (Siddharthan, p. 17) [1]. Text simplification is beneficial for making classic literature available for schoolchildren, for low-literacy communities or people with some types of cognitive impairments.

Over the years, natural language processing (NLP) has evolved significantly, switching from rule-based systems that interpret and generate language to sequence processing. There is a plethora of models, such as ChatGPT, Perplexity or BERT, that use data and artificial intelligence to create human-like language. As this paper is aimed to evaluate ChatGPT's ability to simplify texts, it is relevant to investigate its mechanisms of simplification. Firstly, from a lexical perspective, ChatGPT replaces difficult words with simpler synonyms. The most recent version, ChatGPT 4, recognizes complicated structures and has the capacity to give contextual synonyms. Secondly, it can simplify content on a syntactic level by dividing text into simpler sentences. This enhances coherence and makes the text clearer while maintaining the most important information.

ChatGPT has been used to make texts more available to the public, either specialized or general (Araújo & Aguiar, 2023, Ayre et al., 2023) [2-3]. Many studies prove that AI-generated simplifications improve the quality of higher education by helping students deepen their understanding of topics of their expertise. Furthermore, automated simplifications provided by ChatGPT have proved to be

beneficial in every-day contexts such as: understanding texts related to healthcare (Ayre et al., 2023) [3], cultural events or customer service.

In other words, many studies present ChatGPT's competence to simplify texts, but there is no study designed to assess its simplifications of classic literature taught in schools. Therefore, the purpose of this paper is to investigate the manner in which it simplifies excerpts taken from Mark Twain's *The Adventures of Tom Sawyer*. We chose this novel as it is widely studied in classrooms around the world either translated or in English. While traditional simplified versions of the novel are valuable, this study seeks to explore how AI can further enhance text accessibility, potentially offering more context-sensitive simplifications tailored to individual reader needs. The simplified versions used in our paper are given to primary and middle school students who speak English as second language (L2) who assess the simplification by completing a questionnaire. It is important to mention that this paper is part of a larger study, so we will present only two examples of AI-driven simplifications in order to focus more thoroughly on the process and its outcomes.

2. Literature Review
2.1. ChatGPT: its Relevance in the Field of Text Simplification (TS)

Many language models have emerged in recent years due to the advances made in the domain of natural language processing (NLP). A good example in this sense is the Generative Pretrained Transformer (GPT) that, in turn, encapsulates large language models, such as ChatGPT. It was firstly introduced in December 2022 and since then it has evolved and improved due to what Liu et al. (2023, p. 1) call "the technique of reinforcement learning from human feedback" (RLHF) [4]. The latest version, ChatGPT-4 "empowers exciting developments that involve diverse data beyond text" (Liu et al., p. 1) [4]. This means that this model processes and uses multiple data formats besides text: video, audio, images and many other multimedia types. Many fields, like education, health care, scientific research, economics and so on, have benefited from this expansion by facilitating human-machine interaction, learning and expanding knowledge.

With regard to text simplification and schoolchildren's education, ChatGPT-4 is capable to simplify pieces of text according to the instructions given by the human user. Not only does it shorten sentences, replace difficult words with easier synonyms or omit unnecessary details, but it also provides users with other types of helpful support material that can enhance the understanding of a text. For instance, ChatGPT-4 can make a text easier to be comprehended by turning it into a short clip, or by using visual materials based on the original information. This type of simplification, or better called, transformation can be used as a complementary aid in the education of young learners (middle school level) as it makes lessons more interactive. What is more, in the context of education designed to meet special needs (e.g.: dyslexia, autism spectrum disorder, language processing disorder), visual materials can be used as engaging resources. As Rutherford et al. (2023) claim, visual supports facilitate "receptive or expressive communication" (p. 1) [5], which can significantly improve interaction and learning outcomes for children with communication difficulties.

2.2. Readability Scores

In order for a text to be simplified, its readability score needs to be established so that the initial level of readability is used as a guideline of the simplification. If a text has a low readability score, the simplification does not require many major changes. Nevertheless, if the original text may pose certain problems of understanding, then the criteria of simplification should cover more linguistic aspects.

There are many readability scores that are widely recognized and used, but since this paper is not exhaustive, we will use two of the most important ones, namely the Flesch-Kincaid Grade Level, the Coleman-Liau Index and Coh-Metrix. The former two are relevant for my project even if they are designed for native speakers because both scales provide a general overview of the complexity of texts by measuring aspects like: word difficulty (concerning syllable or character count) and sentence length. Coh-Metrix, on the other hand, is used to evaluate texts for ESL (English as a Second Language) learners too as it analyzes many other layers of the linguistic asset of a written text: lexical diversity, narrative coherence or syntactic complexity.

2.2.2. The Flesch-Kincaid Grade Level

Created in 1975, the Flesch-Kincaid Grade Level is a test that indicates the difficulty of a text by calculating two important aspects: average sentence length and average number of syllables per word (Kincaid et al., 1975) [6]:

*Grade Level = 0.39 * (Total Words / Total Sentences) + 11.8 * (Total Syllables / Total Words) - 15.59*

The final score helps educators, writers or workbooks or content creators to adapt their writings to the target audiences. By determining the readability level of texts, these professionals can make sure that their materials can be understood by readers, enhancing engagement.

2.2.3. The Coleman-Liau Index

The Coleman-Liau Index differs from the Flesch-Kincaid Grade Level in how it measures linguistic complexity: the former uses characters instead of syllables, as the latter does. This makes it easier to be integrated into automatic programs that simplify texts. One of the key advantages is that it can be used in other languages, which makes it more flexible than the Flesch-Kincaid Grade Level. The formula for the index is (Coleman & Liau, 1975) [7]:

CLI=0.0588×L−0.296×S−15.8

(L represents the average number of letters per 100 words; S represent the average number of sentences per 100 words)

2.2.4. Coh-Metrix

Coh-Metrix was develpped in the early 2000s at the University of Memphis and it is different from the other scales as it analyzes text on a cognitive level too (Graesser et al., 2003) [8]. Unlike Flesch-Kincaid and Coleman-Liau, Coh-Metrix goes beyond the superficial surface of the text, by looking into how well texts align with human cognitive capacities and limitations (i.e.: the ability to mentally represent what the text is about; it can measure deep cohesion and that determines how much interfencing is needed so that the readers understand the text, and so on).

3. Methodology
3.1. The Process of Text Simplification

The first step of our analysis was to extract several excerpts from Mark Twain's novel, *The Adventures of Tom Sawyer*. We used University of Oxford's edition from 2007 and we extracted fragments from almost all chapters. Nevertheless, due to the non-exhaustive nature of the article, we will exemplify only two fragments and their simplifications, followed by comments. The extracted fragments were analyzed using readability formulas (Fischer-Kincaid and Coleman-Liau) and an online platform called T.E.R.A. (Coh-Metrix Common Core Text Ease and Readability Assessment) [10]. After setting the original readability scores, we simplified the fragments using the latest version of ChatGPT, namely ChatGPT-4. The simplifications made were mainly syntactical, lexical, but imagery and dialogue (simplification through dialogue) were integrated too. Subsequently, the readability of the simplified texts was assessed using the already-mentioned scales and calculators to show the evolution of the text and its properties.

3.2. The Questionnaire

With the intention to make it more applied, we created a questionnaire using Google Forms, where we compiled the original fragments and their simplifications. The questionnaire was addressed to 21 primary and middle-school children, who answered multiple questions meant to evaluate thier understanding of the simplifed versions.The questionnaire is divided into two sections: the first one dedicated to one excerpt that was simplified from syntactical and lexical perspectives, while the second section contains one excerpt that was simplified through dialogue and imagery.

4. Results

This section is dedicated to analyzing the simplification of each fragment. After analyzing the simplification, we will also interpret the results of the survey addressed to the students.

4.1. The First Fragment

Original fragment: "Tom gave up the brush with reluctance in his face but alacrity in his heart. And while the late steamer "Big Missouri" worked and sweated in the sun, the retired artist sat on a barrel in the shade close by, dangled his legs, munched his apple, and planned the slaughter of more innocents. There was no lack of material; boys happened along every little while; they came to jeer, but remained to whitewash. By the time Ben was fagged out, Tom had traded the next chance to Billy Fisher for a kite, in good repair; and when he played out, Johnny Miller bought in for a dead rat and a string to swing it with— and so on, and so on, hour after hour." (Twain 2007, p. 21)

Simplified fragment: "Tom gave the brush to Ben slowly, but he was happy inside. While the steamboat 'Big Missouri' worked in the sun, Tom sat on a barrel in the shade. He moved his legs, ate an apple, and thought about getting more boys to paint the fence. Many boys came by. They laughed at first but then started to paint. When Ben got tired, Tom traded the brush to Billy Fisher for a kite. When Billy got tired, Johnny Miller traded a dead rat and a string for his turn. This went on for hours." [11]

Simplified fragment (dialogue-based): "Tom gave the brush to Ben slowly, but he was happy inside. While the steamboat 'Big Missouri' worked in the sun, Tom sat on a barrel in the shade. He moved his legs, ate an apple, and thought about getting more boys to paint the fence. Many boys came by. They laughed at first but then started to paint.

"Hey, Ben, you're doing great!" Tom said, munching his apple.

"Thanks, Tom! This is actually kind of fun," Ben replied, wiping sweat from his forehead.

When Ben got tired, Tom traded the brush to Billy Fisher for a kite.

"Billy, you want a turn?" Tom asked.

"Sure, what do you want for it?" Billy responded.

"How about that kite?" Tom suggested.

"Deal!" Billy said, handing over the kite.

When Billy got tired, Johnny Miller traded a dead rat and a string for his turn.

"Hey, Johnny, want to paint?" Tom called out.

"Why not? I'll give you this dead rat and a string," Johnny offered.

"Sounds good to me!" Tom agreed. This went on for hours." [11]

Readability test	Original version	Simplified fragment (1)	Simplified fragment (dialogue-based) (2)
Flesch-Kincaid Grade Level	7.02	3.81	2.14
Coleman-Liau Index	6.38	4.70	3.49
Coh-Metrix (narrativity)	45%	58%	73%
Coh-Metrix (syntactic simplicity)	38%	96%	100%

Table 1. Readability scores of the first fragment

As Table 1 shows, regarding Flesch-Kincaid Grade Level, the original excerpt from *The Adventures of Tom Sawyer* is appropriate for a 7th-grade reading level, while the simplified and dialogue-based versions are suitable for 4th and 2nd graders, respectively. The Coleman-Liau Index also supports this by marking the original fragment with 6.38, and the simplifications with 4.70 and 3.49. The Coh-Metrix analysis made using two criteria, narrativity and syntactic simplicity, highlights that the original version has moderate narrativity, 45%, and syntactic simplicity, 38%. For instance, the original text includes sentences like: "Tom gave up the brush with reluctance in his face but alacrity in his heart," (Twain

2007, p. 21) [9] and "the retired artist sat on a barrel in the shade close by, dangled his legs, munched his apple, and planned the slaughter of more innocents" (Twain 2007, p. 21) [9]. In terms of vocabulary and structure, these two examples are quite difficult to be understood by primary and middle-schoolers.

Conversely, the first simplified and dialogue-based versions show higher narrativity (58% and 73%) and considerably increased syntactic simplicity (96% and 100%). To illustrate this with examples, the simplified version says: "He moved his legs, ate an apple, and thought about getting more boys to paint the fence" [11]. The dialogue-based version further simplifies the text simplification by introducing a short dialogue between Tom and Ben, making it, thus, more accessible by decreasing the mental effort required to understand the text. It also makes it easier for the reader to have a mental representation of the scene, facilitating comprehension. In addition, syntactic complexity is notably reduced in the simplified and dialogue-based texts, making them more accessible.

4.2. The Second Fragment

Original fragment: "A few minutes later Tom was in the shoal water of the bar, wading toward the Illinois shore. Before the depth reached his middle he was half way over; the current would permit no more wading, now, so he struck out confidently to swim the remaining hundred yards. He swam quartering up stream, but still was swept downward rather faster than he had expected. However, he reached the shore finally, and drifted along till he found a low place and drew himself out. He put his hand on his jacket pocket, found his piece of bark safe, and then struck through the woods, following the shore, with streaming garments." (Twain 2007, p. 95)
Simplified fragment: "A few minutes later, Tom was in the low water, walking to the Illinois shore. When the water got deep, he was halfway there. The current was too strong to walk, so he swam the rest of the way. The current pushed him down, but he kept swimming. He got to the shore and found a low spot to get out. He checked his pocket and his piece of wood was safe. Then he walked through the woods with wet clothes." [11]

Readability test	Original fragment	Simplified fragment
Flesch-Kincaid Grade Level	9.51	5.76
Coleman-Liau Index	10.66	7.29
Coh-Metrix (narrativity)	45%	58%
Coh-Metrix (syntactic simplicity)	38%	96%

Table 2. Readability scores of the second fragment

As far as Table 2 is concerned, Coh-Metrix scores reveal significant differences between the original and simplified excerpts. The original excerpt has a Flesch-Kincaid Grade Level of 9.51 and a Coleman-Liau Index of 10.66, and this indicates a higher reading complexity that would be suitable for more advanced readers. In contrast, the simplified version scores 5.76 and 7.29, respectively, which makes it easier to be read by younger audience.

One aspect that is worth mentioning is that the original text uses complex structures and vocabulary, effectively describing Tom's struggle with the water. It includes advanced language, such as: "the current would permit no more wading" (Twain 2007, p. 95) and "he swam quartering up stream, but still was swept downward". These structures emphasize the difficulties Tom encounters while heading to the shore The simplified version, despite being easier to read, does not highlight the struggles, but rather makes the narration more fluid: "The current was too strong to walk, so he swam the rest of the way". So as to compensate for the loss of detail, the simplified text is backed with one collage of three images generated by DALL-E to help children visualize the scene. These images aim to illustrate Tom's struggle with the current and his success in reaching the shore.

4.3. The Outcome of the Questionnaire

As mentioned earlier, the questionnaire was addressed to 21 pupils from primary and middle school and the questions were designed to determine whether the simplifications facilitate understanding or not. As for the question "Are the words easy to understand?", 40.9% and 63.6% opted for "Very easy" when evaluating the first simplification. Interestingly enough, the question "Does the dialogue make the story more engaging and easier to understand?" was answered by 42.9% with "A lot", 52.4% with

"Somewhat" and 4.8% (one response) with "Not really". These percentage indicates that this kind of simplification, through dialogue, may be easier in terms of vocabulary and sentence structure (as the aforementioned tables prove, but that does not necessarily mean that it is suitable for the majority of children. Due to the fact that children have different language levels, some might perceive this kind of simplification more beneficial than others.

In line with the expectations, 90.5% (20 out of 21) answered "Yes" to the question "Do the images make it easier to understand Tom's actions?". This perhaps indicates that visual supports, either images or videos, can enhance children's understanding of the narrated events.

5. Discussions and Conclusion

This study investigates how AI-driven text simplification interacts with readability and comprehension by analyzing several excerpts taken from Mark Twain's *The Adventures of Tom Sawyer*. To provide a clear and focused analysis, we selected two representative fragments for comparison. This allows us to show how simplification made by ChatGPT affects readability metrics and the narrative experience. The original text, with a Flesch-Kincaid Grade Level of 9.51 and a Coleman-Liau Index of 10.66, is suited for older readers with advanced reading skills, while the simplified version, with scores of 5.76 and 7.29 respectively, is more accessible for younger readers with less reading experience. The Coh-Metrix analysis shows that the simplified text increases narrativity and syntactic.

When it comes to text alteration, it is important to no note that ChatGPT's simplied versions lack narrative depth and detailed descriptions as complex vocabulary and details are omitted or reduced. The simplified text, backed with visual aids like DALL-E-generated images, effectively improves comprehension and engagement among readers. This is confirmed by the feedback of the children, who find a text easy to understand and visualize if it is accompanied by visual support. This suggests that visual aids may make literary texts more interesting and appealing, especially for primary-school children. Therefore, we can conclude that ChatGPT's simplifications are well-done and useful when used in the context of adapting texts to certain levels of readability. However, it is relevant to note that AI-driven simplifications come with certain trade-offs, including but not limited to reduction in narrative depth and lack of nuanced descriptions.

REFERENCES

[1] Siddharthan A., "Syntactic simplification and text cohesion", University of Cambridge Computer Laboratory, Cambridge, 2004.
[2] Araújo S., Aguiar M., "Simplifying Specialized Texts with AI: A ChatGPT-Based Learning Scenario", Perspectives and Trends in Education and Technology, Singapore, Springer, 2023, pages 599–609. https://doi.org/10.1007/978-981-99-5414-8_55
[3] Ayre J., Mac O., McCaffery K., et al., "New Frontiers in Health Literacy: Using ChatGPT to Simplify Health Information for People in the Community", J GEN INTERN MED, 2024, pages 573–577. https://doi.org/10.1007/s11606-023-08469-w
[4] Liu Y. J., Han T., Ma S., Zhang J., Yang Y., Tian J., He H., Li A., He M., Liu Z., Wu Z., Zhang L., Zhu D., Li X., Niu Q., Shen D., Liu T., & Ge B., "Summary of ChatGPT-Related research and perspective towards the future of large language models", Meta-radiology, 2023. https://doi.org/10.1016/j.metrad.2023.100017
[5] Rutherford M., Baxter J., Johnston L., Tyagi V., & Maciver D., "Piloting a Home Visual Support Intervention with Families of Autistic Children and Children with Related Needs Aged 0–12", International Journal of Environmental Research and Public Health, 2023. https://doi.org/10.3390/ijerph20054401
[6] Kincaid J. P., Fishburne J., P. R. R., L. C. R., & S. N. B., "Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel", 1975. https://doi.org/10.21236/ada006655
[7] Coleman M., & Liau T. L., "A computer readability formula designed for machine scoring", Journal of Applied Psychology, 1975, pages 283–284. https://doi.org/10.1037/h0076540
[8] Graesser A. C., McNamara D. S., & Louwerse M. M., "What do readers need to learn in order to process coherence relations in narrative and expository text", Rethinking reading comprehension, Guilford Publications, 2003, pages 82-98.

[9]   Twain M., "The Adventures of Tom Sawyer", Oxford University Press, USA, 2007.

[10] McNamara D. S., Graesser A. C., McCarthy P., & Cai Z., "Automated Evaluation of Text and Discourse with Coh-Metrix", Cambridge University Press, Cambridge, 2014.

[11] OpenAI, "ChatGPT (GPT-4)", OpenAI, 2024. https://www.openai.com/chatgpt (Accessed: April 21 - May 17 2024)