Enhancing Textual Accessibility for Readers with Dyslexia through Transfer Learning

Elham Madjidi
Christopher Crick
elham.madjidi@okstate.edu
chriscrick@cs.okstate.edu
Oklahoma State University
Stillwater, OK, USA

ABSTRACT

This paper explores automated modification of text to make it more accessible for people with dyslexia, a reading disorder affecting a significant percentage of the global population. The modifications are both in terms of changing the appearance of text and simplification of the words, grammar, and length of textual material. For simplification of text, we built a dataset with original and dyslexia-friendly text verified by human readers that it improve their reading experience by 27% on average. Then we developed a pipeline to generate dyslexia-friendly text automatically using transfer learning. The model learns styles appropriate for dyslexic users and generates dyslexia-friendly text from arbitrary textual data, which is easier for people with dyslexia to read and interpret.

KEYWORDS

Dyslexic friendly text, style transfer, neural text generation

ACM Reference Format:
Elham Madjidi and Christopher Crick. 2023. Enhancing Textual Accessibility for Readers with Dyslexia through Transfer Learning. In *The 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '23), October 22–25, 2023, New York, NY, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3597638.3614473

1 INTRODUCTION

Dyslexia is a neurological reading disorder marked by problems with accurate and fluent word recognition, as well as poor spelling and decoding skills. Problems with reading comprehension and a lack of reading experience can restrict the development of vocabulary and context information as secondary consequences[11]. This disorder makes it more difficult for people with dyslexia to access written content. Since the vast majority of information is viewed as text, this makes academia extra difficult for them. According to the United Nations, access to information and communication technology should be a basic human right[12].

There are two explanations for approaching textual accessibility for users with dyslexia, both related to its social relevance: first, they are a relatively large community of users, between 5-15% of the global population [2], and second, these kinds of accessibility practices are good not only for people with dyslexia but also for other target groups. As an illustration, text-to-speech (TTS) not only can help people with dyslexia but it is beneficial for a person with visual impairments.

When thinking about modification of text, the first idea is to change how it looks. Then we think of more complicated processes such as reducing the length of the text while preserving the information content, exchanging less frequent words with more familiar ones, and simplifying grammar (for instance, avoiding double negatives).

Even more ambitious is the ability to change the style of a text. Every group of people has their particular writing style, just as styles can be attributed to painters, musicians, and others. Transfer learning, a subfield of artificial intelligence, focuses on separating content from style and learning the style itself. With new content, it can then apply the learned style to generate a new synthesis. This technology forms the foundation of this work, aiming to learn a dyslexia-friendly style and transfer it to new data.

Learning the style of dyslexia-friendly text presented an obstacle: identifying dyslexia-friendly data. To the best of our knowledge, no publicly accessible dataset contains such data or even a curated collection of books and materials specifically generated for people with dyslexia. However, guidelines and suggestions are available for generating dyslexia-friendly textual data. We adopted the most popular of these guidelines and created a set of rules to generate our own dyslexia-friendly text. With our dataset, we developed a machine learning architecture to automatically generate dyslexia-friendly text.

The rest of the article is structured as follows: We present related work in Section 2 that is applicable to our project. The data and model are presented in Sections 3 and 4. The experiment and results is described in Sections 5. Section 6 concludes the article and looks forward to future work.

2 RELATED WORK

In this section, we explore related works for modifying the appearance of text to make it more accessible for people with dyslexia. We also survey works that use transfer learning to generate new data without themselves being subject to the enormous cost of training a model from scratch.

2.1 Text modification

Text presentation and text content matter, according to numerous reports [1, 4, 20, 21]. Readers with dyslexia find that font type has

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
ASSETS '23, October 22–25, 2023, New York, NY, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0220-4/23/10.
https://doi.org/10.1145/3597638.3614473

an impact on readability, as fonts like italics and serif are harder to read and hence should be avoided [17, 18]. Fonts have been created to be dyslexia-friendly[27], but have not shown themselves to be much better than regular fonts [17]. According to one study, font size has a substantial effect on readability; however, line spacing does not [17]. Background colors also play a role in readability, not only for people with dyslexia but for others as well. Warm colors improve readability and cool colors work in the opposite way[18].

The manner in which information is written can have an impact on how easy or difficult it is to understand. Simplification of text can involve various approaches, from replacing complex words with simpler synonyms [19] to condensing lengthy paragraphs into concise sentences. Synonym replacement can be challenging with words that have multiple meanings, and individuals with dyslexia do not typically struggle with understanding complex word meanings, but rather with similar-sounding or similar-looking words [16, 28]. To assist with this, tools such as *SeeWord* can be used to highlight words with similar spellings to draw attention to such differences [29].

2.2 Transfer learning

Transfer learning is a technique used in machine learning to address the challenge of obtaining accurate training data that matches the feature space and data distribution of the test data. This can be a difficult and expensive task, necessitating an alternative solution. Transfer learning leverages knowledge gained from a related source domain to develop a high-performance learner for a target domain. By utilizing a model trained on a different but related problem as a starting point, the learning process becomes more efficient and effective.

Different approaches and techniques exist for transfer learning tasks, categorized based on how information transfer occurs: instances, features, shared parameters of source and target domain learner models, and defined relationships between the domains [26].

Text style transfer aims to automatically manage the style characteristics of written language while preserving the text's content. This technology has diverse potential applications, including the customization of dialogue for individual users in intelligent chatbots, assisting in the creation of more formal or casual language in writing, reducing harmful language in social media posts, and simplifying complex writing [3, 5, 7, 8, 15, 22, 23].

Developing tailored written material for individuals with dyslexia can be a resource-intensive process. Alternatively, text style transfer technology can be utilized to automatically convert standard written material into a more dyslexia-friendly format, providing a potential solution to this challenge.

3 DATA

Although standard written material is widely available on the internet, it can be challenging to find text that is suitable for individuals with dyslexia. In this section, we describe our endeavors to collect and verify (with the help of human assessments) the initial openly accessible dataset of dyslexia-friendly text.

In accordance with the 2018 British Dyslexia Association Style Guide on how to format and present written material [21], we opted to produce dyslexia-friendly versions of sample texts from both the GRE practice exam website [1] and the textbook *Understanding Psychology*[6].

We decided to use these two sources as references due to our university's accessibility center, which allowed us to test the readability of the modified text with college students who have dyslexia. This enabled us to verify whether or not our alterations had actually made the text more easily comprehensible for this population.

To create text that is easier for individuals with dyslexia to read, we utilized the crowdsourcing service Amazon Mechanical Turk [2] (AMT). We provided workers with the guidelines and sample text and asked them to produce modified text that followed those guidelines (Check appendix A.1).

After evaluating the outcomes of the AMT, we selected 3 to 5 altered texts for each of the five passages. Next, we created a survey using Qualtrics[3] that included both the original texts and the modified versions. We then requested students with dyslexia to rate the texts and indicate whether the modified versions were indeed easier to read.

After the initial explanation of the study's objectives, consent form, and demographic inquiries the questionnaire proceeds with five distinct passages and three to five modified versions of each passage. Participants were requested to evaluate each modified version and select one of four categories: no improvement, mild improvement, minor improvement, and huge improvement (check appendix A.2 for a screenshot from the survey). Finally, we asked participants for their overall feedback and suggestions about the survey.

Initially, 75 individuals responded to the survey; however, the number of participants decreased as the survey progressed, and only 7 individuals completed the final question (check appendix appendix:table1 for breakdown). As anticipated, most respondents reported slight improvement for some passages and mild improvement for others. No improvement was a common response for passages with fewer modifications, while huge improvement was the least reported category among all versions (Fig 1).

The report demonstrated that the modifications were beneficial in helping adults with dyslexia read passages with less difficulty. The subsequent phase involves generating additional data similar to the most successful outcomes, developing a dataset, and fine-tuning a model to automatically produce comparable results.

4 MODEL

Automated text production requires advanced neural text generation models, such as transformer-based [24] models like BERT[9], RoBERTa[10], GPT [13], and T5[14]. These models consist of an encoder for the source language and a decoder for the target language, enabling tasks like translation from German to English. Unsupervised pre-trained models, like BERT, learn representations from large unlabelled corpora and can be used across multiple tasks without architectural modifications. By fine-tuning the pre-trained model with an additional output layer for each specific task, these

[1] thttps://www.test-guide.com/free-gre-practice-tests.html
[2] https://www.mturk.com
[3] https://www.qualtrics.com/

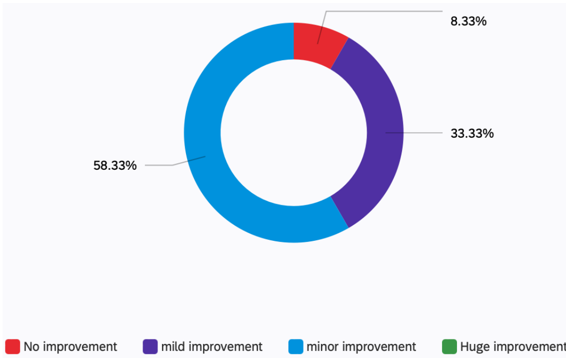

Figure 1: An example of reports for one altered version of one passage. The donut chart has all four colors to show different improvements. The chart shows 58.33% of minor, 33.33%, and 8.33% of no improvement reported by participants for an example passage.

models achieve state-of-the-art performance on various language model tasks, including GLUE[25].

One notable transformer-based language model is the Text-to-Text Transfer Transformer (T5) developed by Google AI Language. T5 has undergone pre-training on a diverse range of text-to-text tasks using a substantial corpus of data. This broad training approach enables T5 to be fine-tuned for tasks like text classification, question answering, summarization, and machine translation, surpassing performance expectations in each area.

To generate dyslexia-friendly text, we utilized two pre-trained models, RoBERTa and T5. The reason for selecting these models was based on their successful performance in tasks similar to our objective. As these large language models already possess comprehensive knowledge of the language, fine-tuning them with dyslexia-friendly text enables them to learn the specific modifications required to generate such text. This is analogous to the approach taken by human workers from AMT, who already have language proficiency but follow the provided guidelines to generate dyslexia-friendly text.

In the next section, we comprehensively talk about the experiments we conducted with models and data.

5 EXPERIMENT

Our experiments aim to investigate the effectiveness of the Roberta and T5 models in generating text similar to our own dataset. We used a dataset of 50 texts with a length of around 100 words per text, which we collected from online sources. The dataset contains the original and changed versions; the changed version is dyslexia-friendly, summarized, and simplified text. The dataset contains a range of topics, including psychology, politics, geology, and others.

For preprocessing, we converted all the text to lowercase and removed any special characters and numbers. We then split the dataset into a training set (80%), a validation set (10%), and a test set (10%).

We fine-tuned the Roberta and T5 models, with a batch size of 4 and a maximum sequence length of 128. We used the Adam optimizer with a learning rate of 2e-5 and a linear learning rate schedule with a warmup of 25 steps. We trained the models for 14 epochs and saved the best checkpoint based on the validation loss.

We evaluated the models on the validation set using the perplexity metric, which measures how well the models predict the probability distribution of the target sequence. We also conducted a human evaluation of the generated text, asking a group of five experts to rate the quality of the text on a scale of 1 to 5, with 5 being the highest quality.

The Roberta and T5 models achieved a perplexity of 31.2 and 25.4, respectively, on the validation set (Fig 2). The human evaluation showed that the Roberta-generated text received an average score of 2.6, while the T5-generated text received an average score of 3.2.

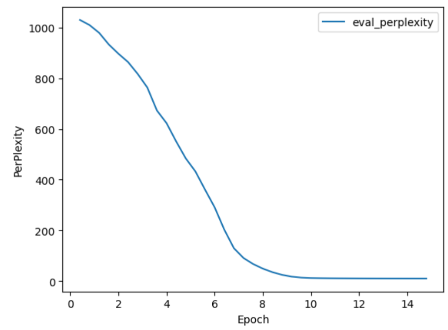

Figure 2: The perplexity of T5 for 14 epochs. An elbow shape curve starts from almost 1000 in the first epoch to the best perplexity in epoch 7 and then goes the same.

Our results suggest that the T5 model is more effective than the Roberta model in generating text similar to our dataset, as it achieved a lower perplexity score and received a higher average score in the typical human evaluation. However, the Roberta model still achieved a reasonable performance.

One limitation of our study is the small size of our dataset, which we plan to increase in future work. The future work will be generating more data on different topics and making another survey with machine-generated text and evaluating it with people with dyslexia.

6 CONCLUSION

We have presented the challenges of dyslexia in accessing textual information and the importance of making information accessible to everyone. We have discussed various text modification techniques that can aid individuals with dyslexia, including font type, size, color, and simplification of text. Additionally, we have explored the potential of transfer learning and text style transfer technology to automatically convert standard written material into a more readable format for individuals with dyslexia. Our efforts to collect and verify a dataset for this purpose have been described in the data section. Overall, this work demonstrates the potential of using machine learning and natural language processing techniques to improve accessibility for individuals with dyslexia.

For future work, we intend to make a more powerful fine-tuned model with a bigger dataset in hand to power an app to scan a page of text from a book and convert it to dyslexia-friendly text while also modifying the appearance of the text based on an individual's preference.

REFERENCES

[1] 2005. The effect of print size on reading speed in dyslexia. *Journal of Research in Reading* 28, 3 (Aug. 2005), 332–349. https://doi.org/10.1111/j.1467-9817.2005.00273.x

[2] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed. ed.). Autor, Washington, DC.

[3] Eleftheria Briakou, Di Lu, Ke Zhang, and Joel Tetreault. 2021. Olá, bonjour, salve! XFORMAL: A benchmark for multilingual formality style transfer. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 3199–3216.

[4] Vagner Figueredo de Santana, Rosimeire de Oliveira, Leonelo Dell Anhol Almeida, and Marcia Ito. 2013. Firefixia: An Accessibility Web Browser Customization Toolbar for People with Dyslexia. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility* (Rio de Janeiro, Brazil) *(W4A '13)*. Association for Computing Machinery, New York, NY, USA, Article 16, 4 pages. https://doi.org/10.1145/2461121.2461137

[5] Cicero dos Santos, Igor Melnyk, and Inkit Padhi. 2018. Fighting Offensive Language on Social Media with Unsupervised Text Style Transfer. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 189–194.

[6] Robert Feldman. 2020. *ISE essentials of understanding psychology* (14 ed.). McGraw-Hill Education, Columbus, OH.

[7] Di Jin, Zhijing Jin, Zhiting Hu, Olga Vechtomova, and Rada Mihalcea. 2022. Deep learning for text style transfer: A survey. *Computational Linguistics* 48, 1 (2022), 155–205.

[8] Di Jin, Zhijing Jin, Joey Tianyi Zhou, Lisa Orii, and Peter Szolovits. 2020. Hooks in the Headline: Learning to Generate Headlines with Controlled Styles. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 5082–5093.

[9] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL-HLT*. 4171–4186.

[10] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* 32 (2019).

[11] International Dyslexia Association: Definition of dyslexia. [n. d.]. Based in the initial definition of the Research Committee of the Orton Dyslexia Society, former name of the IDA, done in 1994. http://interdys.org/DyslexiaDefinition.htm

[12] United Nations Committee of the General Assembly: Convention on the Rights of Persons with Disabilities. 2006. *United Nations, New York* (2006).

[13] Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. 2018. Improving language understanding by generative pre-training. 33 (2018).

[14] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *Journal of Machine Learning Research* 21, 140 (2020), 1–67. http://jmlr.org/papers/v21/20-074.html

[15] Sudha Rao and Joel Tetreault. 2018. Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. 129–140.

[16] Maria Rauschenberger, Luz Rello, Silke Füchsel, and Jörg Thomaschewski. 2016. A language resource of German errors written by children with dyslexia. In *LREC 2016, Tenth International Conference on Language Resources and Evaluation; 2016 May 23-28; Portoroz, Slovenia.[place unknown]: ELRA; 2016. p. 83-7.* ELRA (European Language Resources Association).

[17] Luz Rello and Ricardo Baeza-Yates. 2016. The Effect of Font Type on Screen Readability by People with Dyslexia. *ACM Trans. Access. Comput.* 8, 4, Article 15 (may 2016), 33 pages. https://doi.org/10.1145/2897736

[18] Luz Rello and Ricardo Baeza-Yates. 2017. How to present more readable text for people with dyslexia. *Universal Access in the Information Society* 16 (2017), 29–49.

[19] Luz Rello, Ricardo Baeza-Yates, Stefan Bott, and Horacio Saggion. 2013. Simplify or help? Text simplification strategies for people with dyslexia. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*. 1–10.

[20] Francisca Serrano and Sylvia Defior. 2008. Dyslexia speed problems in a transparent orthography. *Annals of Dyslexia* 58, 1 (2008), 81 – 95. https://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=35395316&site=ehost-live

[21] British Dyslexia Association. Dyslexia style guide. January 2012. *http://www.bdadyslexia.org.uk/* (January 2012).

[22] Minh Tran, Yipeng Zhang, and Mohammad Soleymani. 2020. Towards A Friendly Online Community: An Unsupervised Style Transfer Framework for Profanity Redaction. In *Proceedings of the 28th International Conference on Computational Linguistics*. 2107–2114.

[23] Laurens Van den Bercken, Robert-Jan Sips, and Christoph Lofi. 2019. Evaluating neural text simplification in the medical domain. In *The World Wide Web Conference*. 3286–3292.

[24] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).

[25] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R Bowman. 2019. Glue: A multi-task benchmark and analysis platform for natural language understanding. In *7th International Conference on Learning Representations, ICLR 2019*.

[26] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. 2016. A survey of transfer learning. *Journal of Big data* 3, 1 (2016), 1–40.

[27] Jessica J Wery and Jennifer A Diliberto. 2017. The effect of a specialized dyslexia font, OpenDyslexic, on reading rate and accuracy. *Annals of dyslexia* 67 (2017), 114–127.

[28] MA Wyke. 1986. Reading, Writing and Dyslexia: A Cognitive AnalysisA. W. Ellis, London: Lawrence Earlbaum, 1984, pp. 147. *Behavioural and Cognitive Psychotherapy* 14, 1 (1986), 95–95.

[29] Yeliz Yesilada and Simon Harper. 2019. *Web Accessibility: A Foundation for Research*. Springer.

APPENDICES

A DATASET GATHERING PROCESS

A.1 An example of AMT assignment

The figure below 3 displays a screenshot obtained from the AMT web page, which presents the essential information required by workers to carry out their tasks effectively.

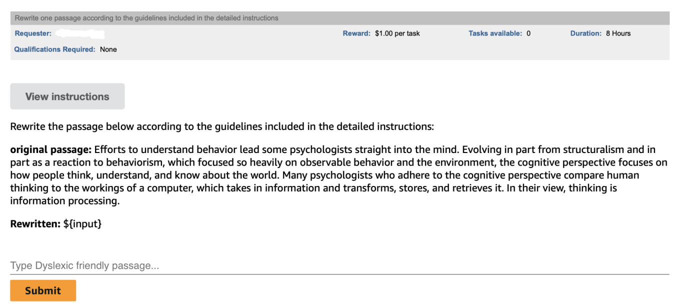

Figure 3: An example of an AMT request. The upper section displays information like the requester's name and reward, while a link directs to the instruction section, which contains specific guidelines for rewriting the passage. The image also includes the original passage, an input bar for workers to type their own rewrite, and a submission button.

A.2 An example survey question

The image 4 represents a survey screenshot, encompassing a paragraph extracted from the GRE exam, along with three distinct versions of texts generated by AMT. Four options are provided for students with dyslexia to select based on their preferences.

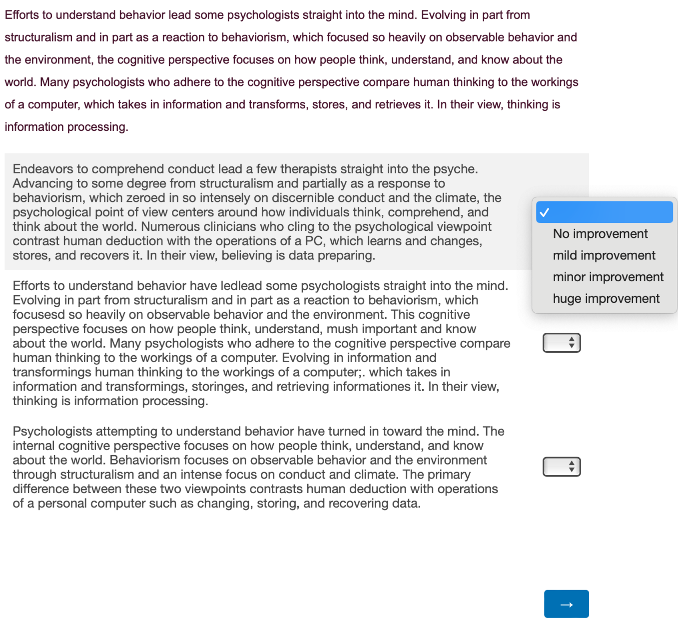

Figure 4: The image depicts an example survey question. At the top, there is a specific question presented. Below it, there are three rewrites of the same passage, each representing a different improvement. As participants progress through each question, the background behind the rewrites turns grey. To select their preferred improvement, participants can click on a scroll button next to each rewrite. Once they have completed all the rewrites, they can proceed to the next passage by clicking the blue arrow button.

A.3 Questions in the survey.

The table presented below (Table 1) displays the count of participants for each survey question, maintaining the same order as they appeared in the survey.

Table 1: Number of participants for each question.

Question	participants
What is your age range?	66
On a scale from 0-10, how severe is your dyslexia?	60
Is English your mother tongue?	63
Passage (from Understanding Psychology)	39
passage (from GRE)	35
Passage (from Understanding Psychology)	35
Passage (from Understanding Psychology)	33
Passage (from GRE)	30
Suggestions and feedback	7