<Keep> it simple!

Controllable Simplified Text Generation in German

Jin Luo, Milena Eisemann, Aydin Javadov

Practical Course on Explainable Artificial Intelligence - WS 22/23

The task of Text Simplification involves rewriting text to make it more legible and comprehensible for individuals with reading difficulties, such as those with cognitive impairments or second-language learners. While the majority of research in the field has been centered on English language, German is investigated rarely, and offers additional challenges as Leichte Sprache follows regulated guidelines. We present a new method for generating Leichte Sprache that additionally enhances control over the generation process via a control token mechanism. This aims to offer greater flexibility in meeting user needs. Finally, we explore different tools that make our methods explainable.

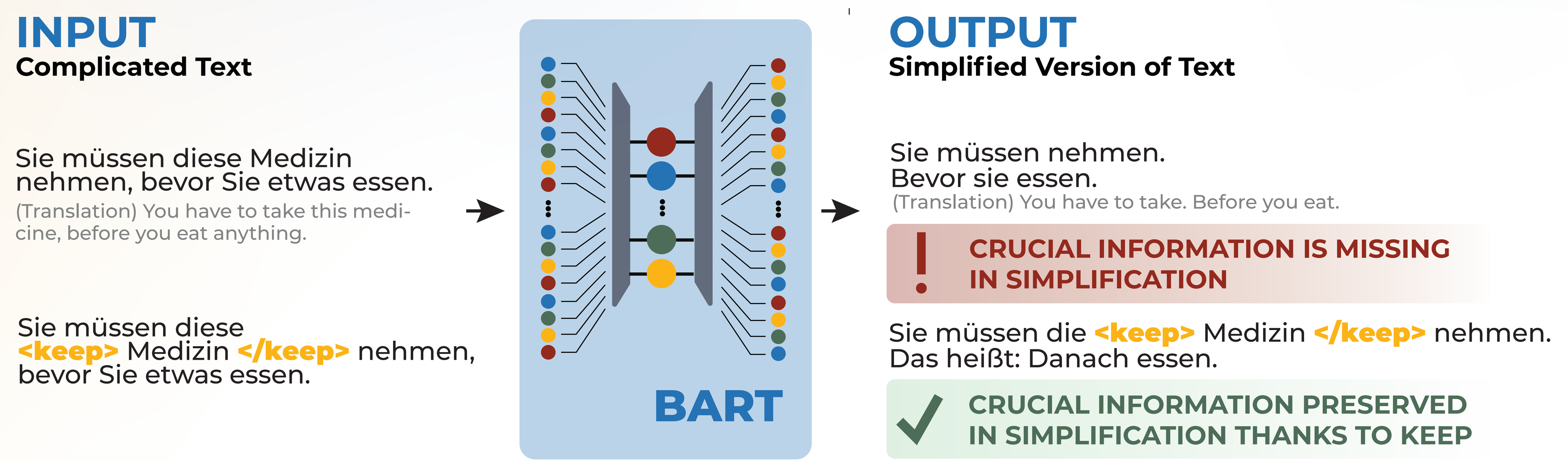


Fig. 1 - Overview of our proposed method for ensuring certain input information is transferred from input to output in a German Leichte Sprache Model

< METHODS >

For our task we fine-tune Bart, T5, mT5, and GPT2 models and subsequently compare their performances. The dataset used in the process consists of 1200 pairs, contains original sentences and their corresponding simplifications, and was provided by SUMM AI GmbH. With the introduction of keep tokens, we aim to ensure that information marked as crucial by the user gets preserved in the simplification (Fig 1). We explore strategies for incorporating these tokens into the training process such as only adding them in the input or in both input and output. Data augmentation is achieved by surrounding one or more words with leading and closing tokens, **<keep>** and **</keep>** respectively. In order to strengthen the ability of the model to learn the <keep> token and reinforce its attention on the enclosed content, we increase the loss based on the keep token part by 10 times and 100 times on the original loss.

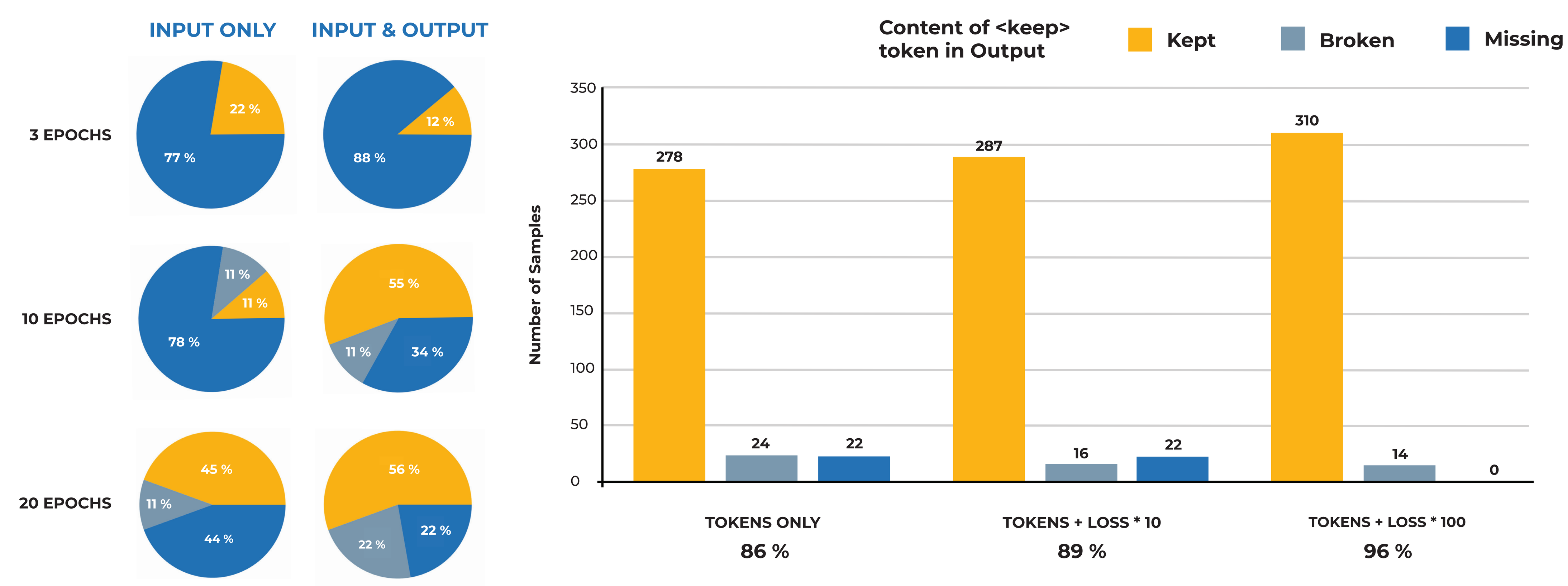
< OUTCOMES >

> Among all models, we identify that Bart generates outputs resembling the format of “Leichte Sprache” the closest. It is hence chosen as our basis model.

> Although longer training stages produce outputs that are more "creative" and "Leichte Sprache"-like, they are frequently poor in grammar. Alternatively, shorter training yields a greater grammar score as the model is merely copying from the input.

> Both basic models (tokens in input only and input & output), are able to influence the outputs to some extent (Fig 2 left). However, their content is not kept reliably enough.

> Using an adapted loss leads to better results in terms of maintaining tokens (Fig 2 right). However, the usage of such enforced loss results in a further decline in grammar.



	No tokens	only input	input & output	loss * 10	loss * 100	reference
Grammar	0.30	**0.60**	0.53	0.34	0.10	-
SARI	35	**42**	35	35	36	-
FRE	57	70	60	69	67	**71**
BERTScore	**0.75**	0.68	0.69	0.73	0.73	-

Fig. 2 - Analysis of amount of correctly kept, partially correct or missing tokens across epochs and models (Top). Comparison of different models using frequently reported metrics for simplification as well as our own grammar score (Bottom).

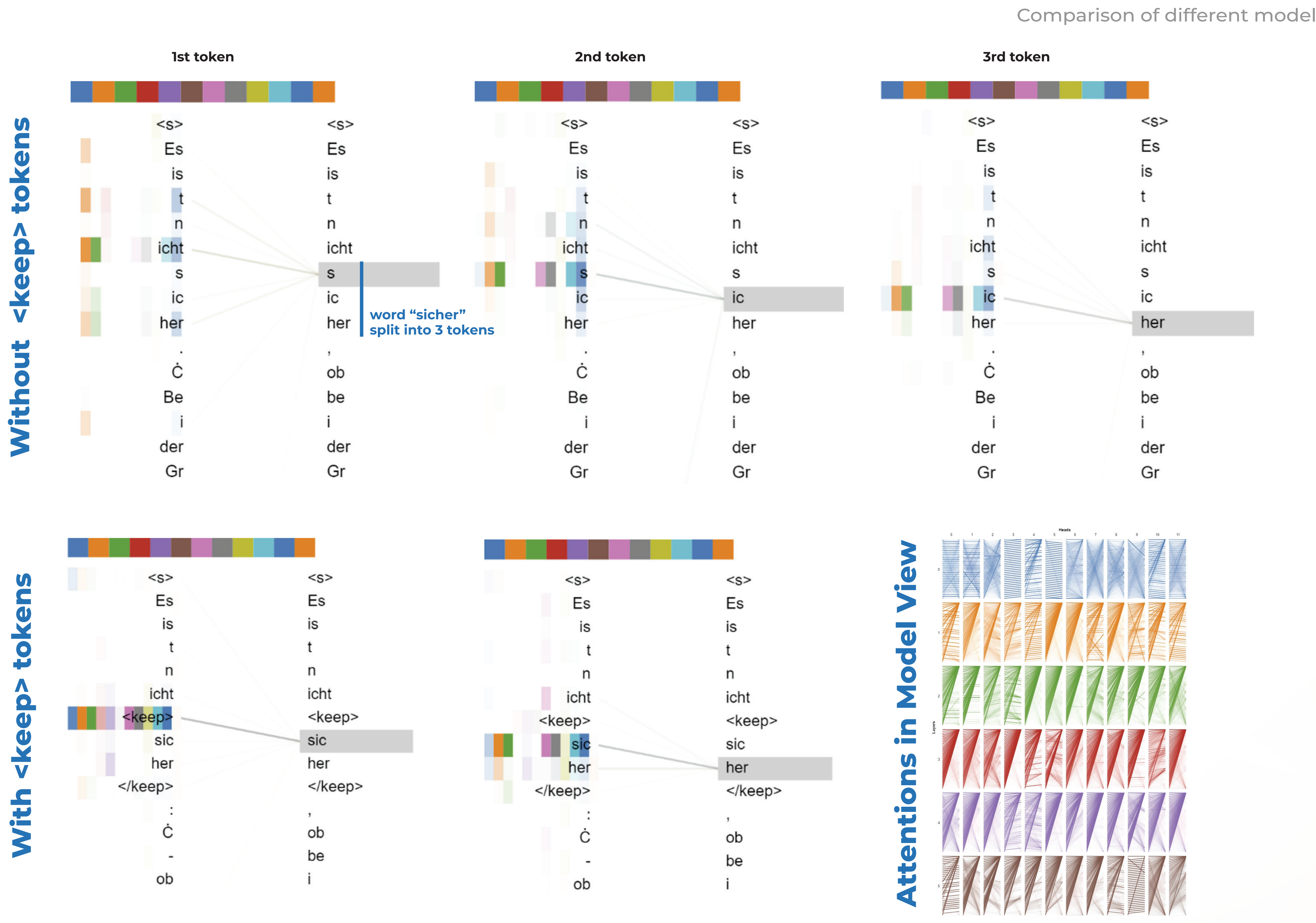


Fig. 3 - Visualizations of cross-attentions in last layer of model with and without tokens, as well as model view. Without tokens, attention is more spread, relating even more distant tokens to current one. <keep> tokens and adapted loss focus attention around the <keep>.

< EXPLAINABILITY >

Comprehending a model's internal workings is crucial to understanding its behavior and decision-making. To achieve this, we delve into the architecture and attention mechanisms of our model. By visualizing both the global model view as well as detailed head views, we are able to examine how the information flows in the model and how the keep tokens influence its predictions (Fig 3).

Furthermore, ease of experimentation is a crucial aspect in the analysis of models. We thus employ a customizable tool that allows for interactive manipulation of the model and data, via features such as word replacements and scrambling as well as various metrics.

< DISCUSSION >

> Efforts are currently restricted by time, ressources and especially the availability of data in German Leichte Sprache.

> Trade-off between creativity (high epochs) and grammar (low epochs).

> <keep> tokens can influence the output to some extend. The adapted loss increases reliability but leads to further decrease in grammar quality. This could be due to the attention being focused on the keep tokens instead of spread across larger parts of the text.

> Interaction via the customized XAI tool is especially valuable for insights into the model’s reasoning and agile experiments.

Future work: Try other surrogate XAI methods (Lime, Anchors), implement other control tokens (e.g. control of output length (4)).

REFERENCES

(1) Advaith Siddharthan. 2014. **A survey of research on text simplification. ITL-International Journal of Applied Linguistics**

(2) Sanja Stajner. 2021. **Automatic text simplification for social good: Progress and challenges**. In Findings of the Association for Computational Linguistics

(3) Jesse Vig. 2019. **A Multiscale Visualization of Attention in the Transformer Model**. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations

(4) Kim Cheng Sheang and Horacio Saggion. 2021. **Controllable Sentence Simplification with a Unified Text-to-Text Transfer Transformer**. In Proceedings of the 14th International Conference on Natural Language Generation