

Differences between German and English Text Simplification

Text Simplification (Alva-Manchego et al., 2020)

Aim

- Automatically reducing the complexity of a given text.
- Generating a more easily understandable version of a complex text.

How to Simplify

- sentence split
- complex word substitution

Evaluation

- Human Evaluation: grammaticality, simplicity, meaning preservation
- Automatic Evaluation: SARI, SAMSA, Flesch Kincaid Grade Level, etc.

Example

Simplification Pair (German):

Complex: "Als Gregor Samsa eines Morgens aus unruhigen Träumen erwachte, fand er sich in seinem Bett zu einem ungeheueren Ungeziefer verwandelt." (Kafka, 1984)

Simple: "Eines Morgens wird Gregor Samsa wach. In der Nacht hat er schon unruhig geträumt. Und als er aufwacht, ist alles anders als sonst. Gregor Samsa ist ein riesiger Käfer geworden. Ein unheimliches Ungezieferes." (Kafka, 2020)

Simplification Pair (English):

Complex: "One morning, as Gregor Samsa was waking up from anxious dreams, he discovered that in bed he had been changed into a monstrous verminous bug." (Kafka, 2003)

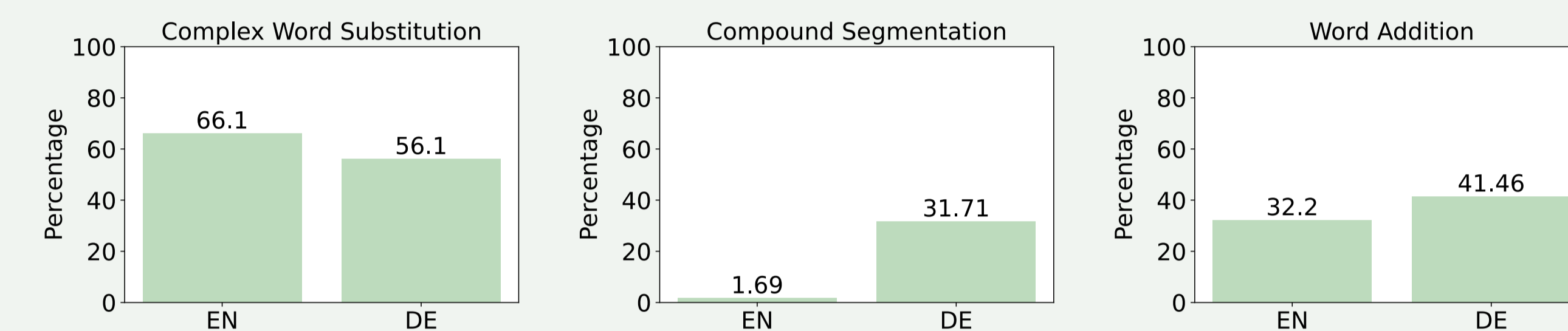
Simple: "One morning Gregor Samsa wakes up. During the night, he has already had restless dreams. And when he wakes up, everything is different than usual. Gregor Samsa has become a giant beetle. A sinister vermin." (own translation)

Plain Language vs. Easy Language

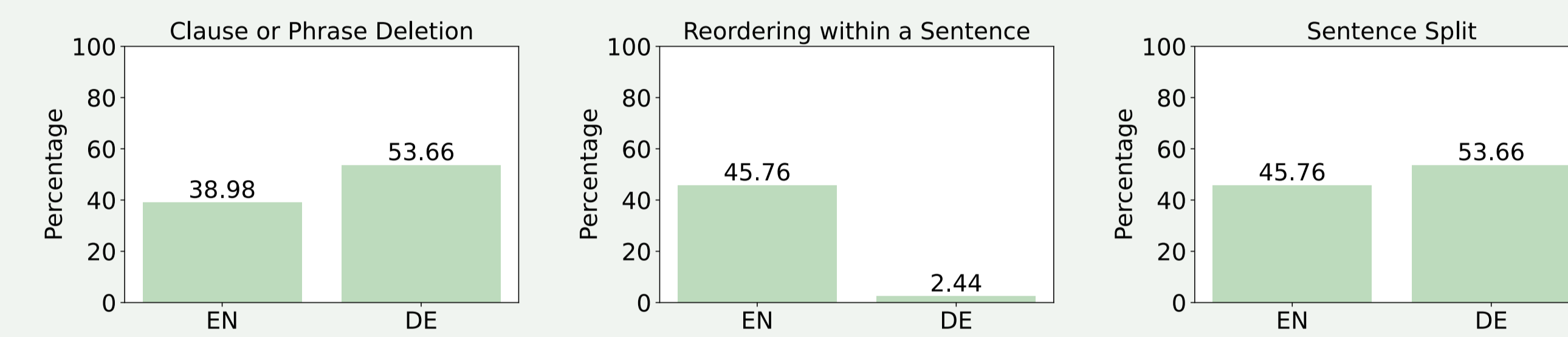
	Plain Language ("Einfache Sprache")	Easy Language ("Leichte Sprache")
Target Group (Reader)	people for whom texts in standard language are too difficult	people with intellectual disabilities and low language skills
Target Group (Writer)	untrained translators	professional translators
Levels of Simplification	word choice, syntax, text composition, style	word choice, syntax, text composition, text design
Characteristics	short but concise simplification	strong contentwise and linguistically reduction
	no unnecessary complication	review of target group required
Aim	simplifying professional jargon for laypeople but keeping the style of speech and generation of proper language	Generating very simplified texts with a large range through strong simplification
References	(Oliveira, 2016), (Baumert, 2018), (Neubauer, 2019)	(Bredel and Maaß, 2018), (Bock, 2019)

Transformations

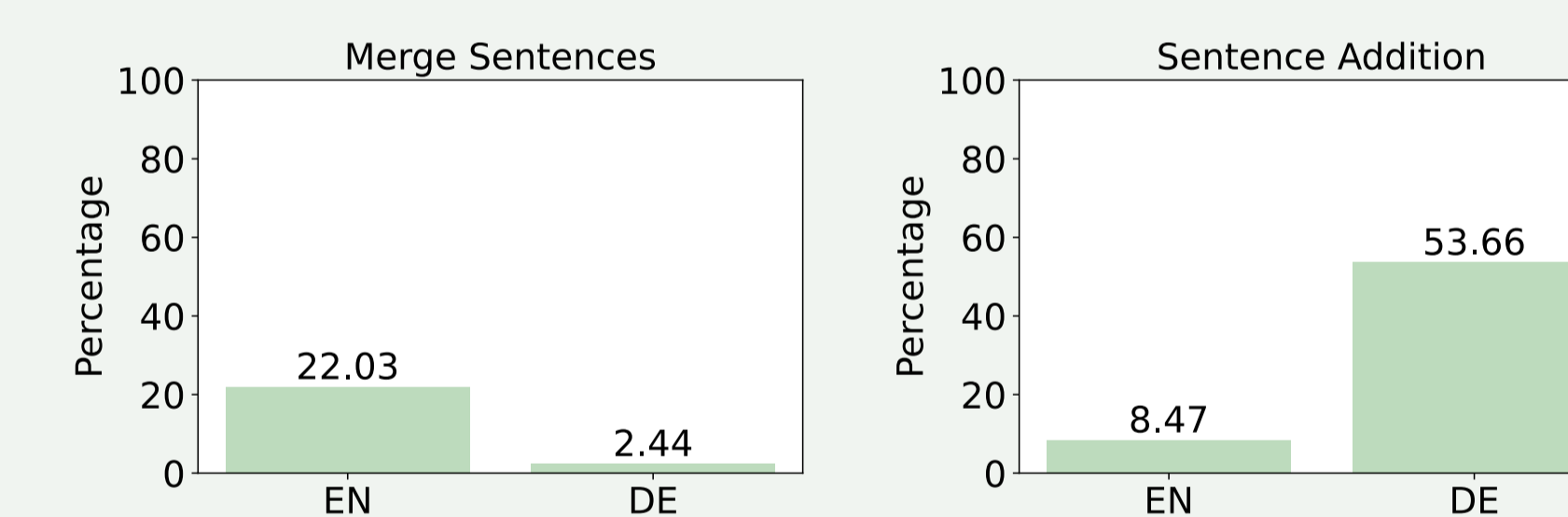
Word Level Transformations



Sentence Level Transformations



Paragraph Level Transformations



Conclusion

- text simplification research is similar to German easy-to-understand language rules
- additional transformations must be considered in German, e.g., addition of sentences or compound segmentation
- German easy-to-understand language is split into plain language and easy language

Future Work

In my dissertation project, I'm working on a text simplification system for German texts. Currently, I'm annotating some German texts with the above named transformations. The number of transformations as well as the realization of the transformations will hopefully lead to insightful information regarding the simplification process. These insights will be included in the text simplification system.

References

Fernando Alva-Manchego, Carolina Scarton, and Lucia Specia. 2020. Data-driven sentence simplification: Survey and benchmark. *Computational Linguistics*, 46(1):135–187.

Andreas Baumert. 2018. *Einfache Sprache: verständliche Texte schreiben*. Spaß am Lesen Verlag, Münster, Germany.

Bettina M. Bock. 2019. "Leichte Sprache" – Kein Regelwerk: Sprachwissenschaftliche Ergebnisse und Praxisempfehlungen aus dem LeISA-Projekt. Frank & Timme, Berlin, Germany.

Ursula Bredel and Christiane Maaß. 2018. *Ratgeber Leichte Sprache: die wichtigsten Regeln und Empfehlungen für die Praxis*. Duden Verlag, Berlin, Germany.

Franz Kafka. 1984. *Die Verwandlung*. Jan Tholenaar Verlag GmbH, Morsbach/Sieg, Münster. Retrieved from Projekt Gutenberg.

Franz Kafka. 2003. *The Metamorphosis*. The Kafka Project. Translated by Ian Johnston of Malaspina University-College, Nanaimo, BC, Canada.

Franz Kafka. 2020. *Die Verwandlung*. Spaß am Lesen Verlag, Münster. Translated by Marion Döbert.

Mansour Neubauer. 2019. *Einfache Sprache - Grundregeln, Beispiele, Übungen*. Eigenverlag (independently published).

Domingos de Oliveira. 2016. *Sagen Sie es einfach: eine Einführung in die einfache Sprache*. Books on Demand, Norderstedt, Germany.