



Automatic Text Simplification for German

Sarah Ebling*, Alessia Battisti, Marek Kostrzewa, Dominik Pfütze, Annette Rios, Andreas Säuberli and Nicolas Spring

Department of Computational Linguistics, University of Zurich, Zurich, Switzerland

The article at hand aggregates the work of our group in automatic processing of simplified German. We present four parallel (standard/simplified German) corpora compiled and curated by our group. We report on the creation of a gold standard of sentence alignments from the four sources for evaluating automatic alignment methods on this gold standard. We show that one of the alignment methods performs best on the majority of the data sources. We used two of our corpora as a basis for the first sentence-based neural machine translation (NMT) approach toward automatic simplification of German. In follow-up work, we extended our model to render it capable of explicitly operating on multiple levels of simplified German. We show that using source-side language level labels improves performance with regard to two evaluation metrics commonly applied to measuring the quality of automatic text simplification.

OPEN ACCESS

Edited by:

Andras Kornai,
Computer and Automation Research
Institute (MTA), Hungary

Reviewed by:

Horacio Saggion,
Pompeu Fabra University, Spain
Julia Fuchs,
Laboratoire STL - Savoirs, Textes et
Langage (UMR 8163), France

*Correspondence:

Sarah Ebling
ebling@cl.uzh.ch

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 07 May 2021

Accepted: 18 January 2022

Published: 23 February 2022

Citation:

Ebling S, Battisti A, Kostrzewa M,
Pfütze D, Rios A, Säuberli A and
Spring N (2022) Automatic Text
Simplification for German.
Front. Commun. 7:706718.
doi: 10.3389/fcomm.2022.706718

Keywords: simplified language, automatic text simplification, automatic sentence alignment, sentence alignment gold standard, German, neural machine translation

1. INTRODUCTION

Simplified language¹ is a variety of standard language characterized by reduced lexical and syntactic complexity, the addition of explanations for difficult concepts, and clearly structured layout. Two tasks deal with automatic processing of simplified language: automatic readability assessment and automatic text simplification (Saggion, 2017).

Automatic text simplification was initiated in the late 1990s (Chandrasekar et al., 1996; Carroll et al., 1998) and since then has been approached by means of rule-based and statistical methods. As part of a rule-based approach, the operations carried out typically include replacing complex lexical and syntactic units with simpler ones (Chandrasekar et al., 1996; Siddharthan, 2002; Gasperin et al., 2010; Bott et al., 2012; Drndarević and Saggion, 2012). A statistical approach (Specia, 2010; Zhu et al., 2010) generally conceptualizes the simplification task as one of converting a standard-language into a simplified-language text using machine translation techniques on a sentence level. The success of such approaches is contingent on the availability of high-quality sentence alignments.

Research on automatic text simplification is comparatively widespread for languages such as English (Zhu et al., 2010), Spanish (Saggion et al., 2015), Portuguese (Aluisio and Gasperin, 2010), French (Brouwers et al., 2014), Italian (Barlacchi and Tonelli, 2013), and other languages. For German, only few contributions exist. Research on simplified German has gained momentum in

¹The term “simplified language” is used to denote the sum of all “comprehensibility-enhanced varieties of natural languages” (Maaß, 2020, p. 52), i.e., what is commonly termed “Easy Language” (German *leichte Sprache*) and “Plain Language” (German *einfache Sprache*). Maaß (2020, p. 52) mentions “easy-to-understand language” as an umbrella term subsuming these varieties. However, in this contribution, we prefer the term “simplified language” to emphasize the notion of the result of a simplification process.

recent years due to a number of legal and political developments in German-speaking countries, such as the introduction of a set of regulations for accessible information technology (*Barrierefreie-Informationstechnik-Verordnung, BITV 2.0*) in Germany, the approval of rules for accessible information and communication (*Barrierefreie Information und Kommunikation, BIK*) in Austria, and the ratification of the United Nations Convention on the Rights of Persons with Disabilities (CRPD) in Germany, Austria, and Switzerland. In addition, two volumes on Easy Language appeared in the “Duden” series (Bredel and Maaß, 2016a,b), further highlighting the relevance of the topic. See Maaß (2020, Chapter 2.3) for a comprehensive overview of the situation in Germany.

The article at hand aggregates the work of our group in automatic processing of simplified German. We present four parallel corpora compiled and curated by our group. We report on the creation of a gold standard of sentence alignments from the four sources for evaluating five alignment methods on this gold standard. We used two of the corpora as a basis for the first sentence-based neural machine translation (NMT) approach toward automatic simplification of German. In follow-up work, we extended our model to render it capable of explicitly operating on multiple levels of simplified German.

More specifically, the contributions of the article at hand are:

- Overview of four parallel (standard/simplified German) corpora, of which automatically generated sentence alignments for one source are available for research purposes
- Gold standard of sentence alignments from the four sources
- Evaluation of automatic sentence alignment methods based on the gold standard
- First sentence-based NMT approach toward automatic simplification of German
- First multi-level simplification approach for German

The remainder of this article is structured as follows: Section 2.1 discusses approaches to automatic sentence alignment in the context of text simplification. Section 2.2 discusses parallel standard-/simplified-language corpora available for language pairs other than standard German/simplified German. Section 2.3 presents previous approaches to automatic text simplification. Sections 3 to 5 present our own contributions, consisting of compiling four standard German/simplified German parallel corpora (Section 3), creating a gold standard for automatic sentence alignment (Section 4) against which to measure existing automatic sentence alignment methods, and performing automatic text simplification on a sentence level (Section 5). Section 6 offers a conclusion and an outlook on future research.

2. PREVIOUS WORK

2.1. Automatic Sentence Alignment

Sentence alignment between standard- and simplified-language texts is an instance of monolingual sentence alignment. As such, it is unable to rely on well-established heuristics of bilingual sentence alignment based on, for example, sentence length (Gale and Church, 1991). The relation between source and target

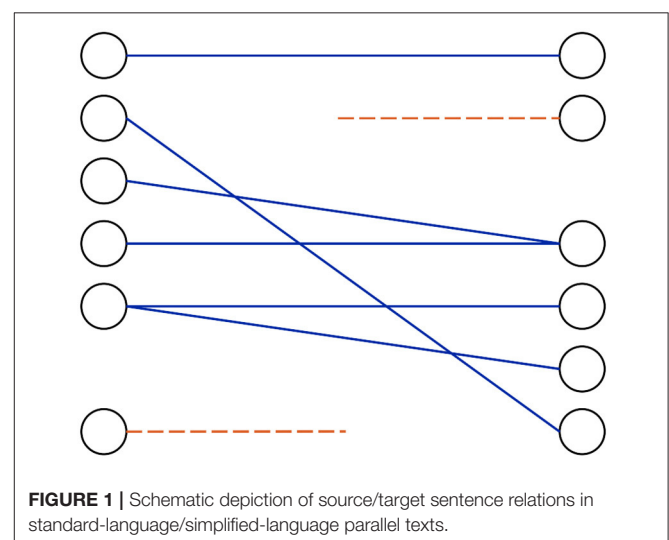
sentences in a standard-language/simplified-language document pair can be of the following types:

- **1:1**, i.e., one standard-language sentence corresponding to one simplified-language sentence
- **n:1** (with $n > 1$), i.e., more than one standard-language sentence reduced to a single simplified-language sentence
- **1:n** (with $n > 1$), i.e., one standard-language sentence split up into multiple simplified-language sentences
- **n:m** (with $n > 1$ and $m > 1$), i.e., more than one standard-language sentence corresponding to more than one simplified-language sentence
- **1:0**, i.e., a standard-language sentence omitted in the simplified-language text
- **0:1**, i.e., a simplified-language sentence inserted compared to the standard-language text

This is visualized in **Figure 1**. Also shown in this figure is an example of a crossing alignment, i.e., an alignment where the order of information of the standard-language text is not the same as that of the simplified-language text (non-monotonicity).

A number of tools have been developed specifically for sentence alignment in the context of text simplification; among them are MASSAlign (Paetzold et al., 2017), CATS (Customized Alignment for Text Simplification) (Štajner et al., 2018), and LHA (Large-scale Hierarchical Alignment for Data-driven Text Rewriting) (Nikolov and Hahnloser, 2019).

MASSAlign is a hierarchical algorithm that uses a vicinity-driven approach. It employs a heuristic according to which the order of information is consistent on the standard- and simplified-language sides, allowing for reduction of the search space. In a first step, MASSAlign searches for alignments between paragraphs, and in a second, for sentence alignments within the aligned paragraphs. The tool employs a similarity matrix with a bag-of-words TF-IDF model with maximum TF-IDF cosine similarity as a similarity metric. The paragraph alignment uses three levels of vicinity: (1) 1:1, 1:n, and n:1 alignments; (2) single-unit skips (where units can be sentences or paragraphs); and (3)



long-distance unit skips. Sentence alignment relies on (1) and (2) only.

Like MASSAlign, CATS is capable of aligning paragraphs and sentences in two steps. The tool offers three similarity strategies, a lexical (character-n-gram-based, CNG) and two semantic similarity strategies. The two semantic similarity strategies, WAVG (Word Average) and CWASA (Continuous Word Alignment-based Similarity Analysis), both require pretrained word embeddings. WAVG averages the word vectors of a paragraph or sentence to obtain the final vector for the respective text unit (sentence or paragraph). CWASA is based on the alignment of continuous words using directed edges. CATS offers two different alignment strategies: MST (Most Similar Text) and MST-LIS (MST with Longest Increasing Sequence) to allow for 1:n alignment.

LHA uses a hierarchical alignment approach with two steps: Firstly, document alignment is performed based on document embeddings and an approximate nearest neighbor search using the Annoy library². Annoy exhibits a low memory footprint *via* usage of static files as indexes. Secondly, sentence embeddings and an inter-sentence similarity matrix are used to extract K nearest neighbors for each source and target sentence. The tool further uses a variation of MST-LIS from CATS to model sentence splitting and compression.

Vecalign (Thompson and Koehn, 2019) and alignment based on SBERT (Reimers and Gurevych, 2020) were introduced in the context of bilingual sentence alignment. SBERT modifies the pretrained BERT network (Devlin et al., 2019) by using siamese and triplet network structures to arrive at sentence embeddings that may then be compared using cosine similarity. Vecalign is a method based on the similarity of the average sentence embedding with cosine similarity as the scoring function.

Table 1 characterizes the five alignment methods MASSAlign, CATS, LHA, SBERT, and Vecalign along the following aspects:

- **All source sentences aligned:** whether the alignment method in its default setup force-aligns every source sentence or bases the decision whether to align a source sentence on a similarity threshold (cutoff)³
- **Concatenation:** whether the alignment method concatenates multiple sentences into one and aligns them as one
- **Crossing alignments:** whether the alignment method allows for abandoning the monotonicity restriction, i.e., supports crossing alignments (cf. **Figure 1**)
- **Alignment type:** which relations between source and target sentences are ultimately supported by the method.

2.2. Sentence-Aligned Parallel Corpora

Automatic text simplification *via* (sentence-based) machine translation as outlined in Section 1 requires pairs of standard-language/simplified-language texts aligned at the sentence level, i.e., parallel corpora. A number of parallel corpora have

²<https://github.com/spotify/annoy> (last accessed: May 5, 2021).

³Note that for CATS, the alignment direction is from simplified language to standard language; hence, CATS searches for one or more standard-language sentences for each simplified-language sentence.

TABLE 1 | Overview of mono- and bilingual sentence alignment tools and methods.

Tool	All source sentences aligned	Concatenation	Crossing alignments	Alignment type
MASSAlign	Cutoff	Yes	No	n:m
CATS	Yes	No	Yes	n:1
LHA	Cutoff	No	Yes	n:m
SBERT	Cutoff	No	Yes	n:1
Vecalign	Yes	Yes	No	n:m

been created to this end. Gasperin et al. (2010) compiled the PorSimples Corpus consisting of Brazilian Portuguese texts (2,116 sentences), each with two different levels of simplifications (“natural” and “strong,”) resulting in around 4,500 aligned sentences. Bott and Saggion (2012) produced the Simplext Corpus consisting of 200 Spanish/simplified Spanish document pairs, amounting to a total of 1,149 (Spanish) and 1,808 (simplified Spanish) sentences (approximately 1,000 aligned sentences).

A large parallel corpus for text simplification is the Parallel Wikipedia Simplification Corpus (PWKP) compiled from parallel articles of the English Wikipedia and the Simple English Wikipedia (Zhu et al., 2010), consisting of about 108,000 sentence pairs. The difference in vocabulary size between the English and the simplified English side of the PWKP Corpus amounts to 18%⁴. Application of the corpus has been criticized for various reasons (Štajner et al., 2018); the most important among these is the fact that Simple English Wikipedia articles are often not translations of articles from the English Wikipedia. Hwang et al. (2015) provided an updated version of the corpus that includes a total of 280,000 full and partial matches between the two Wikipedia versions.

Another frequently used data collection, available for English and Spanish, is the Newsela Corpus (Xu et al., 2015) consisting of 1,130 news articles, each simplified into four school grade levels by professional editors. The difference in vocabulary size between the English side and the simplest level (Simple-4) is 50.8%.

The above-mentioned PorSimples and Newsela corpora present standard-language texts simplified into multiple levels, thus accounting for a recent consensus in the area of simplified-language research, according to which a single level of simplified language is not sufficient; instead, multiple levels are required to account for the heterogeneous target usership.

2.3. Automatic Text Simplification

Specia (2010) introduced statistical machine translation to the automatic text simplification task, using data from a small parallel corpus (roughly 4,500 parallel sentences) for Portuguese. Coster and Kauchak (2011) used the PWKP Corpus in its original form (cf. Section 2.2) to train an MT system. Xu et al. (2016) performed

⁴Vocabulary size as an indicator of lexical richness is generally taken to correlate positively with complexity (Vajjala and Meurers, 2012).

syntax-based MT on the English/simplified English part of the Newsela Corpus (cf. Section 2.2).

Nisioi et al. (2017) pioneered NMT models for text simplification, performing experiments on both the Wikipedia dataset of Hwang et al. (2015) and the Newsela Corpus for English, with automatic alignments derived from CATS (cf. Section 2.1). The authors used LSTMs as instances of Recurrent Neural Networks (RNNs).

More recent contributions to ATS include explicit edit operation modeling (Dong et al., 2019), graded simplification (Nishihara et al., 2019), multi-task learning (Guo et al., 2018; Dmitrieva and Tiedemann, 2021), weakly supervised (Palmero Aprosio et al., 2019), and unsupervised approaches (Surya et al., 2019; Kumar et al., 2020; Laban et al., 2021). These approaches are largely limited to English (Al-Thanyyan and Azmi, 2021) due to a lack of training data in other languages.

Säuberli et al. (2020) presented the first approach to text simplification for German using (sentence-based) NMT models. As data, they used an early version of the APA Corpus (cf. Section 3.2) amounting to approximately 3,500 sentence pairs.

The most commonly applied automatic evaluation metrics for text simplification are BLEU (Papineni et al., 2002) and SARI (Xu et al., 2016). BLEU, the *de-facto* standard metric for machine translation evaluation, computes token n-gram overlap between a hypothesis and one or multiple references. A shortcoming of BLEU with respect to automatic text simplification is that it rewards hypotheses that do not differ from the input. By contrast, SARI was designed to punish such output. It does so by explicitly considering the input and rewarding tokens in the hypothesis that do not occur in the input but in one of the references (addition) and tokens in the input that are retained (copying) or removed (deletion) in both the hypothesis and one of the references. More precisely, SARI computes the arithmetic average of n-gram precision and recall of the three rewrite operations addition, copying, and deletion, specifically rewarding simplifications that are dissimilar from the input. The metric was shown to exhibit “reasonable correlation with human evaluation on the text simplification task” (Xu et al., 2016).

Table 2 displays BLEU and SARI scores for previous sentence-level simplification approaches for different languages.

3. COMPILING DATA FOR AUTOMATIC PROCESSING OF SIMPLIFIED GERMAN

This section reports on our contributions in building and curating four parallel corpora for use in automatic text simplification for German.

3.1. Web Corpus

Klaper et al. (2013) created the first parallel corpus for German/simplified German, consisting of 256 texts each (approximately 70,000 tokens) downloaded from the Web. Battisti et al. (2020) extended the corpus such that it contained more parallel data, newly contained monolingual-only data (simplified German), and newly contained information on text structure (e.g., paragraphs, lines), typography (e.g.,

TABLE 2 | Automatic evaluation scores for sentence-level ATS approaches (PBMT, phrase-based SMT; SBMT, syntax-based MT).

References	Language	Approach	Scores
Specia (2010)	Portuguese	SMT	60.75 BLEU
Coster and Kauchak (2011)	English	SMT	60.46 BLEU
Wubben et al. (2012)		PBMT	67.79 BLEU (Nisioi et al., 2017) 34.07 SARI (Nisioi et al., 2017)
Xu et al. (2016)	English	SBMT	73.62 BLEU (Nisioi et al., 2017) 38.59 SARI (Nisioi et al., 2017)
Nisioi et al. (2017)	English	NMT	87.50 BLEU
Štajner and Nisioi (2018)	English	NMT	Newsela: 89.49 BLEU 36.48 SARI PWKP: 84.69 BLEU 35.78 SARI
Säuberli et al. (2020)	German	NMT	9.75 BLEU 36.88 SARI

font type, font style), and images (content, position, and dimensions)⁵. The parallel part of the corpus is useful for automatic text simplification *via* machine translation (cf. Section 2.3), the monolingual-only part for automatic readability assessment, which is not the focus of this article. In addition, monolingual-only data can also be leveraged as part of machine translation through applying back-translation, a data augmentation technique.

The corpus is compiled from PDFs and webpages collected from Web sources in Germany, Austria, and Switzerland. Information on the underlying guidelines for creating simplified German is not available, as the data was collected automatically. The sources mostly represent websites of governments, specialized institutions, and non-profit organizations. The documents cover a range of topics, such as politics (e.g., instructions for voting), health (e.g., what to do in case of pregnancy), and culture (e.g., introduction to art museums). The corpus contains 6,217 documents, of which 5,461 are monolingual-only, and 378 are available in both standard German and simplified German. The 378 parallel documents amount to 17,121 sentences on the standard German and 21,072 sentences on the simplified German side. Compared to their German counterparts, the simplified German texts in the parallel data have clearly undergone a process of lexical simplification: The vocabulary is smaller by 51% (33,384 vs. 16,352 types), which is comparable to the rate of reduction reported in Section 2.2 for the Newsela Corpus (50.8%).

⁵The importance of the latter type of information has repeatedly been stressed, e.g., for automatic readability assessment (Bredel and Maaß, 2016a; Arfé et al., 2018; Bock, 2018).

TABLE 3 | Examples from the Austria Press Agency (APA) corpus (Säuberli et al., 2020).

Original	<i>15,2 Prozent der Österreicher genießen Gemüse mehrmals am Tag, 41,0 Prozent öfter täglich und der Rest seltener.</i> (‘15.2 percent of Austrians enjoy vegetables several times a day, 41.0 percent often daily and the rest less often.’)
B1	<i>Nur rund 15 Prozent der Österreicher essen mehrmals am Tag Gemüse.</i> (‘Only around 15 percent of Austrians eat vegetables several times a day.’)
Original	<i>Jedes Kalb erhält spätestens sieben Tage nach der Geburt eine eindeutig identifizierbare Lebensnummer, die in Form von Ohrmarken beidseitig eingezogen wird.</i> (‘At the latest seven days after birth, each calf is given a unique identification number, which is recorded on ear tags on both sides.’)
B1	<i>In Österreich bekommt jedes Kalb spätestens 7 Tage nach seiner Geburt eine Nummer, mit der man es erkennen kann.</i> (‘In Austria, at the latest 7 days after birth, each calf receives a number, with which it can be identified.’)
Original	<i>US-Präsident Donald Trump hat in seiner mit Spannung erwarteten Rede zur Lage der Nation seine politischen Prioritäten betont, ohne große wirtschaftliche Initiativen vorzustellen.</i> (‘In his eagerly awaited State of the Union address, U.S. President Donald Trump stressed his political priorities without presenting any major economic initiatives.’)
B1	<i>US-Präsident Donald Trump hat am Dienstag seine Rede zur Lage der Nation gehalten.</i> (‘U.S. President Donald Trump gave his State of the Union address on Tuesday.’)
Original	<i>Sie stehe noch immer jeden Morgen um 6.00 Uhr auf und gehe erst gegen 21.00 Uhr ins Bett, berichtete das Guinness-Buch der Rekorde.</i> (‘She still gets up at 6:00 a.m. every morning and does not go to bed until around 9:00 p.m., the Guinness Book of Records reported.’)
B1	<i>Sie steht auch heute noch jeden Tag um 6 Uhr in der Früh auf und geht um 21 Uhr schlafen.</i> (‘Even today, she still gets up at 6 every morning and goes to bed at 9.’)

3.2. APA Corpus

A second corpus built by our group, which is a parallel corpus throughout, consists of news items of the Austria Press Agency (*Austria Presse Agentur*, APA) with their simplified versions.⁶ At APA, four to six news items per day covering the topics of politics, economy, culture, and sports are manually simplified into two language levels, B1 and A2, following guidelines by *capito*, the largest provider of simplification services (translations and translators’ training) in Austria, Germany, and Switzerland⁷. **Table 3** shows standard German/simplified German (B1) examples from the corpus (Säuberli et al., 2020). The corpus contains a total of 2,426 distinct documents. This amounts to 60,732 standard-language sentences, 30,328 sentences at level B1, and 30,432 sentences at A2. We generated sentence alignments with LHA (cf. Section 2.1), arriving at 10,268 alignments for B1 and 9,456 for A2. The sentence alignments are made available for research purposes⁸.

3.3. Wikipedia Corpus

This parallel corpus was created by automatically translating 150,064 articles of the Simple English Wikipedia (cf. Section

2.2) to German using DeepL⁹ ¹⁰. The synthetically created “simplified German” articles were then aligned on a document level with their standard German counterparts from the German Wikipedia¹¹ using interlanguage links, resulting in 106,126 parallel documents with 6,933,192 standard German sentences and 1,077,992 “simplified German” sentences.

3.4. Capito Corpus

As a provider of simplification services, *capito* produces a high number of professional simplifications for a variety of documents and text genres. This includes but is not limited to booklets, information texts, websites and legal texts, which are manually simplified into one or more levels following the *capito* guidelines. The simplification levels in this corpus include B1, A2, and A1. We extracted simplified German documents along with their standard German counterparts, amounting to 1,055 document pairs for B1, 1,546 for A2, and 839 for A1. The documents contain a total of 183,216 standard-language sentences, 68,529 sentences at level B1, 168,950 sentences at level A2, and 24,243 sentences at level A1. Aligning the sentences with LHA (cf. Section 2.1) yielded 54,224 sentence pairs for B1, 136,582 for A2, and 10,952 for A1.

⁶Note that news items are among the most frequent sources of simplification (Caseli et al., 2009; Klerke and Søgaard, 2012; Bott and Saggion, 2014; Goto et al., 2015; Xu et al., 2015).

⁷<https://www.capito.eu/> (last accessed: August 4, 2020). *capito* distinguishes between three levels along the Common European Framework of Reference for Languages (CEFR) Council of Europe (2009): A1, A2, and B1. Each level is linguistically operationalized, i.e., specified with respect to linguistic constructions permitted or not permitted at the respective level.

⁸<https://zenodo.org/record/5148163> (last accessed: October 14, 2021).

⁹<https://www.deepl.com/translator> (last accessed: May 5, 2021). Simple English Wikipedia authors are instructed to “use Basic English words and shorter sentences”, where Basic English refers to the variety introduced by Ogden (1944) that consists of 850 words on the lexical side.

¹⁰The Simple Wikipedia dump of 12/12/2019 was used, <https://dumps.wikimedia.org/simplewiki/> (last accessed: April 26, 2021).

¹¹Obtained by using the CirrusSearch dump as of 14/09/20, <https://dumps.wikimedia.org/other/cirrussearch/> (last accessed: May 5, 2021).

TABLE 4 | Overview of the four parallel corpora for standard German/simplified German.

Corpus	No. of parallel documents	No. of original sentences	No. of simplified sentences		
			B1	A2	A1
Web Corpus	378	17,121		21,072	
APA Corpus	2,426	60,732	30,328	30,432	n.a.
Wikipedia Corpus	106,126	6,933,192		1,077,992	
Capito Corpus	2,279	183,216	68,529	168,950	24,243

For the Web Corpus and the Wikipedia Corpus, information on language levels is not available.

Table 4 presents an overview of the four data sources.

4. SENTENCE ALIGNMENT GOLD STANDARD AND EVALUATION OF AUTOMATIC SENTENCE ALIGNMENT METHODS

This section reports on the manual creation of a gold standard for sentence alignment based on a subset of the four corpora introduced in Section 3. We subsequently evaluate the five automatic sentence alignment methods presented in Section 2.1 against this gold standard to allow us to select the most accurately aligned sentences as data to train our translation models in Section 5. For more details on this evaluation, see Spring et al. (2021a).

4.1. Method

To create a gold standard against which to measure the performance of the different automatic sentence alignment methods introduced in Section 2.1, we selected approximately 1,500 simplified-language sentences from each of the four sources described in Section 3: the Web Corpus (where 36 documents amount to approximately 1,500 simplified sentences), APA Corpus (134 documents), Wikipedia Corpus (198 documents), and the capito Corpus (42 documents), as summarized in Table 5. Two annotators independently aligned the simplified sentences to their standard-language counterparts, considering all of the alignment types shown in Section 2.1. In case of n:1 or 1:n alignments, the annotators assigned a list of labels of length n to either the standard- or simplified-language sentence. In case of 1:0 or 0:1 alignments, the annotators assigned a placeholder label to the empty standard- or simplified-language sentence. Inter-annotator agreement (Cohen's Kappa) for all corpora was between 0.730 and 0.924 (cf. Table 6). To create a single version of the gold standard, an arbitrator took the final decision in cases where the two annotators disagreed.

4.2. Results

The alignment methods presented in Section 2.1 were used with their default settings and embeddings (where applicable)¹² to

¹²One of the tools, CATS, for example, offers an n-gram-based alignment approach that does not employ embeddings of any kind.

TABLE 5 | Overview of the gold standard of sentence alignments for standard German/simplified German.

Corpus	Parallel documents	Original sentences	Simplified sentences
Web	36	1,454	1,440
APA	134	3,388	1,497
Wikipedia	198	11,668	1,530
capito	42	2,428	1,482
Total	410	18,938	5,949

TABLE 6 | Cohen's Kappa per data source.

Web	Wikipedia	capito A1	capito A2	capito B1	APA A2	APA B1
0.887	0.922	0.924	0.730	0.873	0.885	0.886

TABLE 7 | F1 scores of sentence alignment evaluation from Spring et al. (2021a).

Tool	Web	Wikipedia	Capito A1	Capito A2	Capito B1	APA A2	APA B1
MASSAlign	0.175	0.130	0.096	0.228	0.112	0.076	0.129
LHA	0.339	0.170	0.099	0.321	0.513	0.150	0.213
SBERT	0.218	0.104	0.205	0.348	0.321	0.119	0.136
CATS C3G	0.029	0.037	0.045	0.037	0.032	0.078	0.077
CATS CWASA	0.024	0.035	0.039	0.031	0.024	0.066	0.072
CATS WAVG	0.022	0.031	0.032	0.028	0.026	0.053	0.061
Vecalgn	0.188	0.073	0.215	0.392	0.160	0.085	0.099
Mean	0.142	0.083	0.104	0.198	0.17	0.09	0.112

F1 score represents the mean between precision (how many of the alignments that were extracted by an alignment method were correct in the gold standard) and recall (how many of the alignments in the gold standard were retrieved by an alignment method). Bold: best-performing configuration per column.

align sentences in the pairs of standard-language and simplified-language documents that make up the gold standard. Alignment was performed in both directions, simple to complex and vice versa, and the set of the extracted alignments for both directions was used. This made it possible to evaluate the alignment methods extracting n:1 alignments, even though the gold standard is n:m. Evaluation was performed with the Vecalign scoring script¹³. The scoring script made it possible to evaluate the diverse alignments that naturally occur in text simplification in a standardized way by converting all alignments to a collection of 1:1 alignments.

The results of evaluating the performance of the five alignment methods (MASSAlign, CATS, LHA, SBERT, Vecalign; with CATS featuring three sub-methods) against the gold standard are shown in Table 7 (Spring et al., 2021a). Lower CEFR levels (available in the capito and APA data) proved harder to align and in general corresponded to lower F1 scores. The alignment task becomes harder with increasing distance from standard German, as simplification requires more modifications to the text. Also, on

¹³<https://github.com/thompsonb/vecalign> (last accessed: April 26, 2021).

lower CEFR levels, elaborations and explanations are increasingly common. Generally, the alignment methods performed best on the Web and capito data, with average F1 scores being considerably higher. The low overall scores on the Wikipedia data could be explained by the fact that it is the dataset with the largest disparity between the number of standard German and simplified German sentences (cf. Section 3). Regarding the alignment methods, LHA performed best on five out of the seven datasets. It is also the method with the highest F1 scores on average. On capito A1 and capito A2, Vecalign reached the highest scores.

5. SENTENCE-BASED AUTOMATIC TEXT SIMPLIFICATION

This section reports on our work in training NMT models on two of the data sources introduced in Section 3. For more details, the reader is referred to Spring et al. (2021b).

5.1. Method

For these experiments, we used the APA and the capito corpora introduced in Sections 3.2 and 3.4, respectively, amounting to 19,724 sentence alignments for the APA Corpus (10,268 for B1 and 9,456 for A2) and 201,758 for the capito Corpus (54,224 for B1, 136,582 for A2, and 10,952 for A1), produced with LHA (cf. Section 2.1).

Our baseline models were trained on all available training data across all levels, i.e., these models were language-level-agnostic. They performed generic simplification because they had no explicit method to determine the desired level of simplification on the target side. We trained transformer models (Vaswani et al., 2017) with five layers, four attention heads, 512 hidden units in transformer layers, and 2048 hidden units in transformers feed forward layers. Embedding dropout and label smoothing were set to 0.3. We used BLEU for early stopping on a held-out development set with a patience of 10 checkpoints. We trained with a shared vocabulary (20,000 BPE operations). All our experiments were carried out in sockeye (Hieber et al., 2018).

Our experimental models made use of source-side labels corresponding to the desired CEFR level of the target sentence. These labels allow the model to make a distinction between the different CEFR levels and thus to simplify into different complexity levels. Among others, labels have been used in a variety of tasks such as domain adaption (Kobus et al., 2017), multilingual translation (Johnson et al., 2017), and making better use of back-translation (Caswell et al., 2019). Apart from these modifications to the training data, the model architecture and all hyperparameters were identical to the baseline models and they used the same vocabulary of 20k.

To evaluate our models, we used a test set that consists of 500 parallel sentences each for A1, A2, and B1, which were randomly sampled from the combined corpus.

5.2. Results

The BLEU and SARI scores of our two models on the test sets are presented in **Table 8**. The SARI values of our baseline model are comparable to the results of Säuberli et al. (2020) (cf. **Table 2**), who used a preliminary version of the APA corpus

TABLE 8 | BLEU and SARI scores of the different models.

Model	A1		A2		B1	
	BLEU	SARI	BLEU	SARI	BLEU	SARI
Baseline	13.4	36.26	14.4	36.11	16.3	34.53
APA+capito multi	14.2	43.12	14.1	41.53	17.2	41.81

Bold: best-performing configuration per column.

of approximately 3,500 sentence pairs (cf. Section 2.3), but our baseline achieved higher BLEU scores in the range of 13.4 to 16.3. The experimental model reached improved scores for both metrics. The use of source-side labels boosted performance in terms of BLEU on A1 and B1, with the new values in the range of 14.1 to 17.2. The BLEU score did not improve for A2, which was the level with the highest amount of parallel data available (cf. Section 3). This indicates that the addition of source-side labels may be especially helpful in low resource settings, as, on the other hand, A1 and B1, for both of which there was substantially less data, reached higher scores with the experimental model. In terms of SARI, the addition of source-side labels led to considerable improvements for all levels, with the new scores lying in the range of 41.53 to 43.12.

6. CONCLUSION AND OUTLOOK

This article has presented the work of our group in automatic processing of simplified German. We have given an overview of four parallel corpora compiled and curated by our group: the Web, APA, Wikipedia, and capito corpora. Moreover, we have reported on the creation of a gold standard of sentence alignments from the four sources for evaluating five alignment methods on this gold standard (MASSAlign, CATS, LHA, SBERT, Vecalign; with CATS featuring three sub-methods). We found that LHA performed best on five out of the seven datasets (Web, Wikipedia, capito A1, capito A2, capito B1, APA A2, APA B1). It was also the method with the highest average F1 scores (on capito A1 and capito A2, Vecalign reached the highest absolute scores). In general, for the multi-level sources (capito and APA), lower CEFR levels proved harder to align and corresponded to lower F1 scores. Intuitively, the alignment task becomes harder with increasing distance from standard German, as simplification requires more modifications to the text. Also, on lower CEFR levels, elaborations and explanations are increasingly common. Generally, the alignment methods performed best on the Web and capito data, with average F1 scores being considerably higher. The low overall scores on the Wikipedia data can be explained by the fact that it is the dataset with the largest disparity between the number of standard German and simplified German sentences.

We used the LHA alignments as a basis for the first sentence-based neural NMT approach toward automatic simplification of German (baseline model), and we proposed a model that is capable of explicitly operating on multiple levels of simplified German. We showed that compared to our baseline model, this multi-level experimental model reached improved scores for both automatic evaluation metrics, BLEU and SARI. Specifically,

performance improved on all levels with respect to SARI and on A1 and B1 with respect to BLEU (A2 is the level with the highest amount of parallel data available).

We plan to further investigate the potential of the various alignment methods by varying the embedding strategies and the cutoff values used. In doing so, we expect to further increase the performance of our text simplification approaches according to automatic metrics. In addition, we plan to evaluate the output of future models with the help of human experts and to investigate the comprehensibility of the output among the target groups, e.g., persons with cognitive impairments.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because data from the commercial provider of simplification services (capito) is not publishable. Sentence alignments based on APA have, however, been made publicly available here: <https://zenodo.org/record/5148163>.

AUTHOR CONTRIBUTIONS

SE as the group leader was involved in the creation of the parallel corpora, the sentence alignment gold standard, and

the text simplification experiments. AS, NS, and AR carried out the text simplification experiments. AB was the primary person responsible for the Web corpus and one of the annotators of the sentence alignment gold standard. DP was the second annotator. NS was the arbitrator and performed the evaluation of the sentence alignment methods relative to the gold standard. MK provided the sentence alignments. All authors contributed to the article and approved the submitted version.

FUNDING

This submission received funding through the *capito automatisiert* project funded by the Austrian Research Promotion Agency (*Österreichische Forschungsförderungsgesellschaft*, FFG) General Programme under grant agreement number 881202.

ACKNOWLEDGMENTS

The authors are greatly indebted to the Austria Presse Agentur (APA) and CFS GmbH for providing data consisting of standard German documents with their simplified counterparts.

REFERENCES

- Al-Thanyyan, S. S., and Azmi, A. M. (2021). Automated text simplification: a survey. *ACM Comput. Surveys (CSUR)* 54, 1–36. doi: 10.1145/3442695
- Aluisio, S. M., and Gasperin, C. (2010). “Fostering digital inclusion and accessibility: the porsimples project for simplification of portuguese texts,” in *Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas* (Los Angeles, CA), 46–53.
- Arfé, B., Mason, L., and Fajardo, I. (2018). Simplifying informational text structure for struggling readers. *Read. Writ.* 31, 2191–2210. doi: 10.1007/s11145-017-9785-6
- Barlacchi, G. and Tonelli, S. (2013). “ERNESTA: a sentence simplification tool for children’s stories in Italian,” in *Proceedings of the 14th Conference on Intelligent Text Processing and Computational Linguistics (CICLing)* (Samos), 476–487.
- Battisti, A., Pfütze, D., Säuberli, A., Kostrzewa, M., and Ebling, S. (2020). “A corpus for automatic readability assessment and text simplification of german,” in *Proceedings of The 12th Language Resources and Evaluation Conference* (Marseille: European Language Resources Association), 3295–3304.
- Bock, B. (2018). “*Leichte Sprache*–Kein Regelwerk. Sprachwissenschaftliche Ergebnisse und Praxisempfehlungen aus dem LeiSA-Projekt. Technical Report, Universität Leipzig.
- Bott, S., and Saggion, H. (2012). “Automatic simplification of Spanish text for e-Accessibility,” in *Proceedings of the 13th International Conference on Computers Helping People with Special Needs (ICCHP)* (Linz), 527–534.
- Bott, S., and Saggion, H. (2014). Text simplification resources for Spanish. *Lang. Resour. Eval.* 48, 93–120. doi: 10.1007/s10579-014-9265-4
- Bott, S., Saggion, H., and Figueroa, D. (2012). “A hybrid system for Spanish text simplification,” in *Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies* (Montreal, QC), 75–84.
- Bredel, U. and Maaß, C. (2016a). *Leichte Sprache: Theoretische Grundlagen. Orientierung für die Praxis* (Berlin: Duden).
- Bredel, U., and Maaß, C. (2016b). *Ratgeber Leichte Sprache: Die wichtigsten Regeln und Empfehlungen für die Praxis* (Berlin: Duden).
- Brouwers, L., Bernhard, D., Ligozat, A.-L., and François, T. (2014). “Syntactic sentence simplification for French,” in *Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)* (Gothenburg) 47–56.
- Carroll, J., Minnen, G., Canning, Y., Devlin, S., and Tait, J. (1998). “Practical simplification of English newspaper text to assist aphasic readers,” in *Proceedings of the AAAI’98 Workshop on Integrating AI and Assistive Technology* (Madison, Wisconsin), 7–10.
- Caseli, H. M., Pereira, T. F., Specia, L., Pardo, T. A. S., Gasperin, C., and Aluisio, S. M. (2009). “Building a Brazilian Portuguese parallel corpus of original and simplified texts,” in *10th Conference on Intelligent Text Processing and Computational Linguistics* (Mexico City), 59–70.
- Caswell, I., Chelba, C., and Grangier, D. (2019). “Tagged back-translation,” in *Proceedings of the Fourth Conference on Machine Translation (Volume 1: Research Papers)* (Florence: Association for Computational Linguistics), 53–63.
- Chandrasekar, R., Doran, C., and Srinivas, B. (1996). “Motivations and methods for text simplification,” in *Proceedings of the 16th Conference on Computational Linguistics* (Copenhagen), 1041–1044.
- Coster, W., and Kauchak, D. (2011). “Learning to simplify sentences using wikipedia,” in *Proceedings of the Workshop on Monolingual Text-To-Text Generation (MTTG)* (Portland, OR), 1–9.
- Council of Europe (2009). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Cambridge: Cambridge University Press).
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). “BERT: pre-training of deep bidirectional transformers for language understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (Minneapolis, MN: Association for Computational Linguistics), 4171–4186.
- Dmitrieva, A., and Tiedemann, J. (2021). “Creating an aligned Russian text simplification dataset from language learner data,” in *Proceedings of the 8th Workshop on Balto-Slavic Natural Language Processing* (Kiyv: Association for Computational Linguistics), 73–79.
- Dong, Y., Li, Z., Rezagholizadeh, M., and Cheung, J. C. K. (2019). “EditNTS: an neural programmer-interpreter model for sentence simplification through explicit editing,” in *Proceedings of the 57th Annual Meeting of the Association*

- for *Computational Linguistics* (Florence: Association for Computational Linguistics), 3393–3402.
- Drndarević, B., and Saggion, H. (2012). “Towards automatic lexical simplification in Spanish: an empirical study,” in *Proceedings of the First Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)* (Montreal, QC), 8–16.
- Gale, W. A., and Church, K. W. (1991). “A program for aligning sentences in bilingual corpora,” in *29th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference* (Berkeley, CA), 177–184.
- Gasparin, C., Maziero, E., and Aluisio, S. M. (2010). “Challenging choices for text simplification,” in *Computational Processing of the Portuguese Language. Proceedings of the 9th International Conference, PROPOR 2010* (Porto Alegre), 40–50.
- Goto, I., Tanaka, H., and Kumano, T. (2015). “Japanese news simplification: task design, data set construction, and analysis of simplified text,” in *Proceedings of MT Summit XV* (Miami, FL), 17–31.
- Guo, H., Pasunuru, R., and Bansal, M. (2018). “Dynamic multi-level multi-task learning for sentence simplification,” in *Proceedings of the 27th International Conference on Computational Linguistics* (Santa Fe, NM: Association for Computational Linguistics), 462–476.
- Hieber, F., Domhan, T., Denkowski, M., Vilar, D., Sokolov, A., Clifton, A., et al. (2018). “The sockeye neural machine translation toolkit at AMTA 2018,” in *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (AMTA)*, Vol. 1. 200–207.
- Hwang, W., Hajishirzi, H., Ostendorf, M., and Wu, W. (2015). “Aligning sentences from standard wikipedia to simple wikipedia,” in *Proceedings of NAACL-HLT* (Denver, CO), 211–217.
- Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., et al. (2017). Google’s multilingual neural machine translation system: enabling zero-shot translation. *Trans. Assoc. Comput. Linguist.* 5, 339–351. doi: 10.1162/tacl_a_00065
- Klaper, D., Ebling, S., and Volk, M. (2013). “Building a German/simple german parallel corpus for automatic text simplification,” in *ACL Workshop on Predicting and Improving Text Readability for Target Reader Populations* (Sofia), 11–19.
- Klerke, S., and Søgaard, A. (2012). “DSim, a Danish parallel corpus for text simplification,” in *Proceedings of Language Resources and Evaluation Conference (LREC)* (Istanbul), 4015–4018.
- Kobus, C., Crego, J., and Senellart, J. (2017). “Domain control for neural machine translation,” in *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017* (Varna: INCOMA Ltd.), 372–378.
- Kumar, D., Mou, L., Golab, L., and Vechtomova, O. (2020). “Iterative edit-based unsupervised sentence simplification,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5–10, 2020*, eds D. Jurafsky, J. Chai, N. Schluter, and J. R. Tetreault (Association for Computational Linguistics), 7918–7928.
- Laban, P., Schnabel, T., Bennett, P., and Hearst, M. A. (2021). “Keep it simple: unsupervised simplification of multi-paragraph text,” in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (Association for Computational Linguistics), 6365–6378.
- Maaß, C. (2020). *Easy Language – Plain Language – Easy Language Plus. Balancing Comprehensibility and Acceptability Easy – Plain – Accessible*, Vol. 3 (Berlin: Frank & Timme).
- Nikolov, N. I., and Hahnloser, R. (2019). “Large-scale hierarchical alignment for data-driven text rewriting,” in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)* (Varna: INCOMA Ltd.), 844–853.
- Nishihara, D., Kajiura, T., and Arase, Y. (2019). “Controllable text simplification with lexical constraint loss,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (Florence: Association for Computational Linguistics), 260–266.
- Nisioi, S., Štajner, S., Ponzetto, S. P., and Dinu, L. P. (2017). “Exploring neural text simplification models,” in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Vancouver, BC), 85–91.
- Ogden, C. K. (1944). *Basic English: A General Introduction With Rules and Grammar*. Vol. 29. ed K. Paul (London, UK: Trench, Trubner).
- Paetzold, G., Alva-Manchego, F., and Specia, L. (2017). “Massalign: alignment and annotation of comparable documents,” in *Proceedings of the IJCNLP 2017, Tapei, Taiwan, November 27 - December 1, 2017, System Demonstrations*, eds S. Park and T. Supnithi (Tapei: Association for Computational Linguistics) 1–4.
- Palmero Aprosio, A., Tonelli, S., Turchi, M., Negri, M., and Di Gangi, M. A. (2019). “Neural text simplification in low-resource conditions using weak supervision,” in *Proceedings of the Workshop on Methods for Optimizing and Evaluating Neural Language Generation* (Minneapolis, MI: Association for Computational Linguistics), 37–44.
- Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). “BLEU: a method for automatic evaluation of machine translation,” in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)* (Philadelphia, PA), 311–318.
- Reimers, N., and Gurevych, I. (2020). “Making monolingual sentence embeddings multilingual using knowledge distillation,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Punta Cana: Association for Computational Linguistics), 4512–4525.
- Saggion, H. (2017). *Automatic Text Simplification*. (San Rafael, CA: Morgan & Claypool Publishers).
- Saggion, H., Štajner, S., Bott, S., Mille, S., Rello, L., and Drndarević, B. (2015). “Making it simple: implementation and evaluation of a text simplification system for Spanish,” *ACM Trans. Accessible Comput. (TACCESS)* 6:14.
- Säuberli, A., Ebling, S., and Volk, M. (2020). Benchmarking data-driven automatic text simplification for German. “In *Proceedings of the 1st Workshop on Tools and Resources to Empower People with READING Difficulties (READI)* (Marseille: European Language Resources Association), 41–48.
- Siddharthan, A. (2002). “An architecture for a text simplification system,” in *Proceedings of the Language Engineering Conference* (Hyderabad), 64–71.
- Specia, L. (2010). “Translating from complex to simplified sentences,” in *Computational Processing of the Portuguese Language. Proceedings of the 9th International Conference, PROPOR 2010* (Porto Alegre), 30–39.
- Spring, N., Pfützte, D., Kostrzewa, M., Battisti, A., Rios, A., and Ebling, S. (2021a). “Comparing sentence alignment methods for automatic simplification of german texts,” in *Presentation Given at the 1st International Easy Language Day Conference (IELD)*, (Germersheim).
- Spring, N., Rios, A., and Ebling, S. (2021b). “Exploring German multi-level text simplification,” in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP)*, 1339–1349.
- Štajner, S., Franco-Salvador, M., Rosso, P., and Ponzetto, S. P. (2018). “CATS: a tool for customized alignment of text simplification corpora,” in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Miyazaki: European Language Resources Association (ELRA)).
- Štajner, S., and Nisioi, S. (2018). “A detailed evaluation of neural sequence-to-sequence models for in-domain and cross-domain text simplification,” in *Proceedings of the 11th Language Resources and Evaluation Conference* (Miyazaki).
- Surya, S., Mishra, A., Laha, A., Jain, P., and Sankaranarayanan, K. (2019). “Unsupervised neural text simplification,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (Florence), 2058–2068.
- Thompson, B., and Koehn, P. (2019). “Vecalign: improved sentence alignment in linear time and space,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (Hong Kong: Association for Computational Linguistics), 1342–1348.
- Vajjala, S., and Meurers, D. (2012). “On improving the accuracy of readability classification using insights from second language acquisition,” in *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP (NAACL HLT ’12)* (Montreal, QC), 163–173.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). “Attention is all you need,” in *Advances in Neural Information Processing Systems 30*, eds I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.), 5998–6008.
- Wubben, S., van den Bosch, A., and Krahmer, E. (2012). “Sentence simplification by monolingual machine translation,” in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (Jeju Island: Association for Computational Linguistics), 1015–1024.

- Xu, W., Callison-Burch, C., and Napoles, C. (2015). Problems in current text simplification research: new data can help. *Trans. Assoc. Comput. Linguist.* 3, 283–297. doi: 10.1162/tacl_a_00139
- Xu, W., Napoles, C., Pavlick, E., Chen, Q., and Callison-Burch, C. (2016). Optimizing statistical machine translation for text simplification. *Trans. Assoc. Comput. Linguist.* 4, 401–415. doi: 10.1162/tacl_a_00107
- Zhu, Z., Bernhard, D., and Gurevych, I. (2010). “A monolingual tree-based translation model for sentence simplification,” in *Proceedings of the International Conference on Computational Linguistics* (Beijing), 1353–1361.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ebling, Battisti, Kostrzewa, Pfütze, Rios, Säuberli and Spring. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.