Text Simplification in the Field of Environment and Climate Change: Exploring ChatGPT for Scientific Popularisation in LSP Classes

Katrin Herget[1], Teresa Alegre[2]

University of Aveiro | Centre for Languages, Literatures and Cultures, Portugal[1]
University of Aveiro | Centre for Languages, Literatures and Cultures, Portugal[2]

Abstract

Due to the importance of Languages for Specific Purposes (LSP) [1] in our globalised society, Higher Education institutions have to prepare students for the acquisition of domain-specific knowledge in different field areas. The exploration of specialised corpora assumes a key role in that respect, as it provides access to domain-specific texts that allow the study and description of linguistic patterns and phenomena of a certain area. The use of comparable corpora as an important auxiliary for the analysis and description of LSP provides the students with valuable tools for data-driven learning (DDL) [2]. Working with domain-specific texts demands, apart from the identification of specific linguistic patterns, the ability to differentiate among a variety of communicative situations. These communicative situations require competence in the application of specific registers that may range between specific domain language and popular discourse. The aim of this study is to investigate to what extent ChatGPT can be used to simplify specialized texts. This paper intends to enhance students' competencies in assessing machine-generated simplified text segments in German and Portuguese from the field of environment and climate change policies. The quality assessment is based on a specific set of different categories that were taken and adapted from the Multidimensional Quality Model (MQM) framework. The application of MQM criteria for post-editing simplified ChatGPT text provides students with essential skills that help them improve their language awareness in handling machine-generated content.

Keywords: *Languages for Specific Purposes (LSP), domain-specific texts, machine-generated text simplification, environment and climate change policies, ChatGPT, German-Portuguese.*

1. Introduction

This small-scale pilot study explores the benefits and limitations of using ChatGPT in teaching LSP. The aim is to study to what extent ChatGPT is suitable and adequate for text simplification of specialized texts. This paper intends to enhance students' competencies in assessing machine-generated simplified text segments in German and Portuguese from the field of environment and climate change policies, thus contributing to human evaluation in an academic context. The quality assessment is based on a specific set of different categories that were taken and adapted from the MQM framework. The application of MQM criteria for human annotation of simplified text produced by ChatGPT provides students with essential skills that help them improve their linguistic and cultural awareness in regard to machine-generated content.

In an era marked by swift globalization, the need for expertise in multiple languages across various fields has increased significantly. This has necessitated the instruction of language specialists with specialized knowledge in their respective domains [3] (p. 35).

In this context, multilingual communication has assumed a crucial role in industries that demand language skills for specific purposes, enabling seamless cooperation and information exchange in diverse linguistic environments. As the professional landscape becomes progressively intricate and professions continue to diversify, individuals are increasingly required to specialize according to evolving requirements. Recently, the emergence of artificial intelligence (AI) has brought about a new paradigm in language teaching and learning: The creation of Large Language Models like ChatGPT, which has led to extensive research on its potentialities and limitations for language teaching and learning [4], [5].

2. LSP and ChatGPT

According to Scarpa (2020) [6], special languages are "language varieties found in documents with a

predominant emphasis on the information they convey and directed to a more or less restricted target specialist community, ranging from experts to laypersons and having very specific professionally or subject-related communicative needs and expectations" (p. 3). From a wider perspective, LSP refer to the use of language for academic and professional interactions within specialized domains, such as science, technology, and business and "can be applied to education offered on any language or to multilingual education" [7] (p. 101). LSP are characterized by "distinctive terminological features" and "a specialised use of textual, syntactical and lexical features" [6] (p.3)

In the 21st century the integration of technology in LSP classes in higher education has transformed language education, enhancing interactive learning and providing students with valuable digital skills for a specific domain. The technological upswing has significantly reshaped LSP by granting access to specialized discourse and communication through ICT [8]. Recently, the advent of AI-powered intelligent chatbots has sparked new discussions and debates in the field of language teaching and learning. Koehnke et al. [4] explore pedagogical benefits of using ChatGPT for language learning. The improvement in the field of Large Language Models like ChatGPT has made them increasingly well-suited for various natural language-related tasks, such as simplification, summarization and translation. The objective of this study centres around the topic of text simplification.

3. Text simplification

It is commonly accepted that text simplification is a highly complex activity, as it is carried out at various levels – lexical, syntactical, textual and pragmatical. Extensive research has been conducted in the field of text simplification, with a strong focus on two primary methodologies for assessment: automated metrics and human evaluation metrics. These two approaches serve as key pillars in evaluating the effectiveness and quality of text simplification techniques. Automatic text simplification is a natural language processing (NLP) technique that aims at making spoken or written content more accessible. According to Saggion [9] it is "the process of transforming a text into another text which, ideally conveying the same message, will be easier to read and understand by a broader audience"(p.xiii).

In our research, we apply the human evaluation methodology within the context of LSP teaching. Our objective is to equip students with the skills required for becoming proficient editors, thereby enhancing their ability to handle linguistic and cultural aspects.

4. Method and study design

For assessing the quality of texts simplified by ChatGPT, we employed a Language Quality Assessment (LQA) model that can help measure the machine-generated output, and contribute to a more objective analysis. Such models aim to provide insights into how well machine-generated texts produce valuable results in terms of fluency, coherence, relevance, and other linguistic attributes.

For the present study, we applied the MQM (Multidimensional Quality Model) Framework, which is well-suited for human annotation of machine output, allowing for a multilevel evaluation of text quality across a set of different dimensions or criteria. Each dimension corresponds to a specific aspect of quality, such as accuracy, fluency, style, terminology, etc. The MQM framework was developed in response to the EU-funded QTLaunchPad and QT21 projects by Lommel et al. [10] as a framework for evaluating machine translation quality. Although text simplification usually implies the assessment of machine-generated output at intralingual level, we applied the MQM framework to provide students with competencies in the evaluation of AI generated simplified text. For our pilot study, we selected a limited set of MQM criteria that we consider adequate for the assessment of simplified texts.

The following three categories were applied (Table 1).

Categories	Subcategories	Description
Linguistic Convention	grammar	- Errors in verb tenses, modes, prepositions, etc., affecting grammatical correctness.
	punctuation	- Missing or incorrect usage of punctuation marks, impacting sentence structure and clarity.
	spelling	- Spelling mistakes that can reduce overall readability.

Categories	Subcategories	Description
	language variant	- Mixing different language variants (e.g., European Portuguese and Brazilian Portuguese), causing inconsistency.
Simplicity	not simplified	- text maintains high degree of technicality, making it less accessible for the general reader
	unnecessary simplification	- words/phrases from general language are overly simplified.
Accuracy	addition	- Inclusion of incorrect information, altering the intended meaning of a word/phrase.
	misinterpretation	- Alteration of the meaning of a segment due to the selection of incorrect words or structures.
	omission	- Omission of relevant information, potentially leading to incomplete or misleading content.
Style	awkward	- Usage of inappropriate complex sentence structures.
	inconsistent	- Inconsistent use of formal and informal language or lack of cohesion.

The scientific articles in the German language were collected from the website of the publisher Springer, which allows the search for scientific articles by field, such as Earth Sciences, Environment, and subfields, such as Ecotoxicology, Environmental Health. Part of the articles is made available in open access. The scientific articles from the mentioned website are available in five languages: English, German, Dutch, French, and Italian. Regarding the compilation of Portuguese scientific articles, the absence of a Portuguese language counterpart to the Springer publishing house imposed a different approach. Initially, a search was conducted on Google Scholar, employing specific search terms like "environmental sciences" (ciências ambientais) and "climate change" (alterações climáticas). Subsequently, scientific articles identified within specialized academic journals, such as the Revista de Recursos Hídricos / Journal of Water Resources, were systematically gathered. Furthermore, the abstracts in Portuguese language of these articles were extracted and gathered in a corpus. For the purpose of this pilot study, we compiled two domain-specific comparable text corpora in German and Portuguese for text simplification, each consisting of 30 original text segments and their simplified equivalents, the latter generated in ChatGPT. The segmentation of both corpora allowed for a more focused analysis.

5. Results and discussion
In the following, we present and discuss errors found within the analyzed categories in both the German and Portuguese corpora, providing examples from both the original segments (OS) and their corresponding simplified segments (SS).
5.1 Linguistic Convention
In terms of Linguistic Convention, the evaluation of the German corpus, revealed errors in subcategory Grammar, such as an incorrect use of prepositions, which can be observed in segment 1.

(OS1DE) (...) was zum Ersatz sensibler <u>durch</u> tolerante Arten führen kann.
(SS1DE) Das kann dazu führen, dass empfindliche Arten *<u>von</u> robusteren Arten ersetzt werden.
(Korrektur) Das kann dazu führen, dass empfindliche Arten <u>durch</u> robustere Arten ersetzt werden.

For the same subcategory, we also observed the incorrect conversion of passive voice to active voice. In German scientific texts passive verb structures are typical, as they convey information in a more objective and formal manner.

(OS2DE) Die Struktur terrestrischer Lebensgemeinschaften <u>wird</u> durch anthropogene Beeinträchtigung <u>verändert</u>.
(SS2DE) *Doch <u>menschliche Aktivitäten verändern</u>, wie diese Lebensgemeinschaften funktionieren.

In the Portuguese corpus, similar errors to the German corpus could be found in regard to the choice of preposition, or active/passive voice. Example 1 illustrates the choice of active instead of passive voice, which confers the wrong idea of a personal action. Additionally, this example reveals an error of grammatical number "the temperature" instead of "the temperatures".

(OS1PT) <u>As temperaturas</u> de superfície <u>foram extraídas</u> da imagem Landsat

(SS1PT) *<u>Tiramos a temperatura</u> da superfície da imagem do satélite Landsat

(Correction) <u>As temperaturas</u> de superfície <u>foram medidas a partir</u> da imagem do satélite Landsat

The mixture of the language variants European Portuguese and Brazilian Portuguese occurred consistently along the corpus. This is a relevant fact, because in the instruction given to ChatGPT is was explicitly mentioned that the simplification should be verbalized in the European variant. In Portugal, the common designation for en. *researcher* is *investigador* and not *pesquisador*.

(OS2PT) A boa comunicação e dinâmica colaborativa entre <u>investigadores</u>, stakeholders e decisores é também fundamental

(SS2PT) É crucial que *<u>pesquisadores</u>, partes interessadas e tomadores de decisão trabalhem juntos

5.2 Simplicity

With regard to the second category, Simplicity, the German text segment analysis revealed two types of errors: a) overly simplified text, and b) omitted simplification. The first category means simplifying words or phrases that do not require simplification, resulting in a potential loss of content and misrepresentation through an overly reduced complexity.

(OS3DE) <u>die Kenntnisse</u> zur Entstehung, Konzentrationsvariabilität und Wirkung von Ozon kurz und kritisch bewertend zusammenzufassen

(SS3DE) *<u>die Informationen</u> zur Entstehung, Schwankungen in der Konzentration und die Wirkung von Ozon kurz und kritisch zu bewerten

In other cases, ChatGPT missed opportunities to simplify words or phrases, which could enhance accessibility. For example, adjectives such as *global* or *heterogen* could have been replaced with another German adjective.

(OS4DE) Darüberhinaus ist die Ozonbelastung zu einem <u>globalen</u> Problem geworden.

(SS4DE) Zusätzlich ist die Ozonbelastung zu einem globalen Problem geworden.

(OS5DE) (…) inbesondere <u>heterogene</u> Prozesse

(SS5DE) insbesondere durch *<u>heterogene</u> Prozesse

In the Portuguese segments, most of the identified errors belonged to the first category. In example 5, the deictic *montante* (upstream) and *jusante* (dowstream) are unnecessarily reformulated by the phrases *parte de cima* (upper side) and *parte de baixo* (underside), which are confusing for the reader in relation to a river description. A possible simplification, if needed, would be *do lado da nascente* (headwaters) and *do lado da foz* (rivermouth).

(OS3PT) três localizados no rio Mondego (<u>montante</u>, intermédio, <u>jusante</u>)

(SS3PT) três no próprio rio (*<u>parte de cima</u>, parte do meio e *<u>parte de baixo</u>)

Nevertheless, we could also identify a word - *biodiversidade* (biodiversity) - in a segment that would be clearer for the lay audience if it would have been paraphrased by v*ariedade das diferentes espécies de animais e vegetais*.

(OS4PT) a preservação da <u>biodiversidade</u> destes organismos nos ecossistemas aquáticos está diretamente relacionada com a conservação e preservação dos cursos de água em que habitam

(SS4PT) proteger a *<u>biodiversidade</u> desses seres significa cuidar dos rios onde vivem

5.3 Accuracy

The third category, Accuracy, relates to semantics and implies a change in the meaning of a word or phrase, which can impact the factual correctness and integrity of the conveyed information.

(OS5DE) das industriell geprägte Zentrum Hamburgs

(SS5DE) *das Industriezentrum Hamburg

In the Portuguese corpus, we identified two categories of errors: Lack of accuracy by misinterpretation (example 5) and by omission (example 6). In example 5, we notice that "several" (*vários*) has changed to "many" (*muitos*) in the simplified segment, which alters the original information.

(OS5PT) <u>Vários</u> autores destacam a importância dos macroinvertebrados bentónicos

(SS5PT) *<u>Muitos</u> especialistas dizem que os pequenos seres vivos que vivem no fundo dos rios

In example 6, the omission of relevant information in the simplified segment prevents the full comprehension of the message. In the original text, it it said that this is kind of information can help "decision making" (*adequada à tomada de decisões*), whereas the simplified segment only states that "better information" (*melhores informações*) can be obtained.

(OS6PT) disponibilidade de <u>informação adequada à tomada de decisão</u>

(SS6PT) obter <u>informações melhores</u>

5.4 Style

The fourth and last category, Style, refers to inadequate complex sentence structures and lack of cohesion. In the Portuguese corpus, we found text segments in which the demonstrative pronoun *isso* ("that") did not refer to the corresponding message.

(OS7PT) A gestão da água em cenários de alterações climáticas é uma oportunidade <u>para rever metodologias, criando soluções inovadoras</u>

(SS7PT) Lidar com a gestão da água em tempos de mudanças climáticas é uma <u>chance de encontrar novas maneiras de fazer *isso</u>
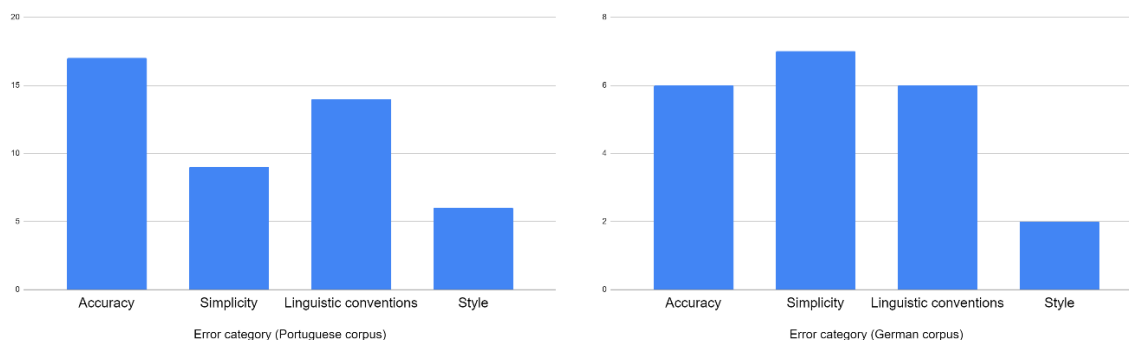

Table 2 - Results of error analysis in the Portuguese and German corpora

6. Conclusions

The results of this small-scale pilot study show that ChatGPT provides assistance in simplifying domain-specific texts from the field of environment and climate change. AI generated text simplification can serve as an informational resource about a specialized text, providing basic information on a text. However, the output analysis showed that human annotation is still an essential prerequisite for high-quality text simplification. The obtained error list of the examined language pair clearly indicates that human error annotation is indispensable and must be carried out carefully, in order to assess the suitability of the AI generated text across multiple categories. Human text simplification involves the expertise of individuals who understand the target audience and the context of the content. When instructing students on how to apply a set of categories to evaluate a machine-generated simplified text, we aimed at equipping them with fundamental tools for linguistic and cultural text comprehension. This empowers them to apply their world knowledge effectively in the process of quality assessment, thus developing their cognitive abilities at various linguistic levels.

References

[1]   Kastberg, P. "Knowledge Communication: Prolegomenon to a Research Programme", *Communication & Language at Work* Vol. 8(1), 2022, pp. 33-49.
[2]   Gilquin, G., & Granger, S. "How can data-driven learning be used in language teaching?", *The Routledge Handbook of Corpus Linguistics*, 2010, pp.359-370.
[3]   Gollin-Kies, S., Hall, D.R., & Moore, S. *Language for Specific Purposes.* Palgrave Macmillan, 2016.
[4]   Kohnke, N., Moorhouse, B.L., & Zou, D."Using Chat GPT for Language Teaching and Learning. *RELC Journal*, 54(2), 2023, 537-550. https://doi.org/10.1177/00336882231162868.
[5]   Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F. et al. "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. EdArXiv. https://doi:10.35542/osf.io/5er8f.https://doi.org/10.35542/osf.io/5er8f
[6]   Scarpa, F. *Research and Professional Practice in Specialised Translation.* Palgrave Studies in Translating and Interpreting, Palgrave-Macmillan, 2020.
[7]   Koskela, M, & Isohella, S. "Teaching Lsp To Technical Communicators: An International Handbook", *Languages for Special Purposes*, De Gruyter, 2018, pp. 96-110.
[8]   Arnó-Macià & Mancho-Barés. "The role of content and language in content and language integrated learning (CLIL) at university: Challenges and implications for ESP", English for specific Purposes 37, 63-73
[9]   Saggion, H. *Automatic text simplification,* Morgan & Claypool, 2017.
[10] Lommel, R., Burchardt, A., & Uszkoreit, H. Multidimensional quality metrics: a flexible system for assessing translation quality. In *Proceedings of Translating and the Computer 35*, 2013.