

Translating Easy Language administrative texts: a quantitative analysis of DeepL’s performance from German into Italian using a bilingual corpus

Christiane Maaß

University of Hildesheim / Germany
maass@uni-hildesheim.de

Chiara Fioravanti

National Research Council of Italy
Institute of Legal Informatics and Judicial Systems / Italy
chiara.fioravanti@cnr.it

Abstract

This study evaluates the performance of DeepL as an AI-based translation engine, in translating German Easy Language Texts into Italian. The evaluation is quantitative and based on a corpus of 26 German fact sheets and their Italian human translations. The results show that DeepL’s translations exhibit significant errors in terminology, accuracy, and language conventions. The machine-translated texts often lack consistency in terminology, and the use of technical or unfamiliar words is not adapted to the difficulty level of the target language. Furthermore, the translations tend to normalize the texts towards standard administrative language, making them less accessible. The study highlights the need for human post-editing to ensure both accuracy and suitability of the translated texts. The findings of this study will help identify where to prioritize post-editing efforts and facilitate comparisons with the results obtained from other artificial intelligence tools used for interlingual translation of Easy Language texts in the administrative domain.

1 Introduction

Easy Language, a comprehensibility-optimized form of a natural language that makes content accessible to people with communication

impairments (Maaß, 2015; Bredel & Maaß, 2016; Maaß, 2020; Maaß & Schwengber, 2022), can play an important role in institutional communication, enabling greater civic participation and inclusion. However, the extent to which it is adopted for legal and administrative texts is not the same from an international perspective (Lindholm & Vanhatalo, 2021a), leading to very different amounts of texts that are available for the different European languages. In German-speaking countries, for example, like Germany, Switzerland, and Austria, the use of Easy Language in public communication is a common and well-established approach (see Maaß et al., 2021; Parpan-Blaser et al., 2021; Fröhlich & Candussi, 2021). In this perspective, interlingual translation could be a valuable asset in expanding the use of Easy Language, all the more so with AI tools at hand.

In a previous study (Maaß & Fioravanti, in press) we examined the feasibility of utilizing DeepL, an AI-based translation engine, recognized for its high accuracy (Fitria, 2023; Kaplan, 2021), as a machine translation tool for interlingual translation into Easy Language within the domain of administrative communication for the language pair German and Italian. The performance analysis of DeepL was based on a corpus derived from texts in Easy Language produced, both in German and Italian, by the administration of the Province of Bolzano/Bozen (a multilingual geographical area in Italy).

In this study, we quantify the errors present in the machine-translated Italian target texts in comparison with the gold standard human translations.

2 Related Work

2.1 Easy Language

Easy Language (also referred to as Easy-to-read, on terminological issues see Lindholm & Vanhatalo, 2021b, and the contributions for the various languages in the *Handbook of Easy Languages in Europe* (Lindholm & Vanhatalo, 2021a)) is a comprehensibility-optimized version of a natural language (for German, see the work of Maaß, 2020; Bredel & Maaß, 2016; for Italian, see the work of Sciumbata, 2022, and Perego, 2021). Vanhatalo & Lindholm (2021a) describe the situation of Easy Language for 20 European countries. In doing so, they not only look at regulations and the legal situation, but also at the text types and domains for which Easy Language texts are available for the various languages. A particularly large number of legal-administrative communication texts are available for German (Rink, 2020; Maaß et al., 2021; Maaß & Rink, 2021). It is therefore reasonable to make these texts usable for other languages via interlingual translation. Particularly in the case of European legal topics or in multilingual regions and communities (Ahrens & Fioravanti, 2022) an increase in the number of available texts for the various languages involved can be expected.

It is also important to acknowledge that legal and administrative texts require a significant effort for translation into Easy Language due to their specialized terminology, complex syntax, and reliance on knowledge of legal procedures (Maaß & Rink, 2021). It is logical, then, to adapt an established best practice across different languages. This approach was implemented in the Province of Bolzano/Bozen, where legal and administrative texts were first translated intralingually into Easy German and subsequently into Easy Italian.

2.2 Machine Translation into or between Easy Languages

Recently, there has been growing interest in exploring machine translation in the context of Easy Language and Plain Language. However, the focus has always been on intralingual translation (see, for example, the work of Deilen et al., 2023, Deilen et al., 2024a, Deilen et al., 2024b). This is obvious, as most translations into Easy Language and Plain Language are intralingual (Maaß, 2020;

Maaß, 2024). However, Pedrini (2024) shows that interlingual translation into Plain Language is also a common practice. There is significant research desideratum here. In a previous study (Maaß & Fioravanti, in press), the authors of this paper have already explored the possibilities of interlingual machine translation between the language pair German-Italian (both directions).

2.3 Evaluating the quality of translations via MQM

In the present paper, the Multidimensional Quality Metrics (MQM) framework was employed. MQM offers a comprehensive catalog of more than 100 issue types that encompass all key translation quality assessment metrics. These issues serve as a "master catalog" from which the most relevant metrics can be selected to evaluate specific translation quality tasks. As an open and freely available framework, MQM can be adopted and expanded to suit various needs (Lommel et al., 2014).

From the MQM CORE Typology error the following four categories were chosen in line with the approach of Ahrens et al. (in press), who have analyzed errors in machine translation of simplified texts: Terminology, Accuracy, Language conventions and Audience appropriateness. However, Ahrens et al. (in press) focus on intralingual translation.

3 Research Design

The analysis of DeepL's performance was based on a corpus extracted from texts in Easy Language from the Province of Bolzano/Bozen in Italy, a bilingual region where both German and Italian are official languages

These texts were produced, in German and Italian, through a collaboration between the provincial administration and Okay, the Easy Language Office of the non-profit organization Lebenshilfe ("live aid"). The Easy language texts are available on the official website of the Province of Bolzano, in a dedicated section (<https://lingua-facile.provincia.bz.it/>).

The German Easy Language texts were created following the rules established by the Research Centre for Easy Language at the University of Hildesheim, as outlined by Maaß (2015) and Bredel & Maaß (2016). They were proof-read by readers with intellectual disabilities (on Easy Language for this target group in Germany see

Maaß & Maaß, 2024). The Italian Easy Language texts originate from the translation of the German versions while also incorporating specific guidelines for Italian Easy Language, as defined by Sciumbata (2022). Like the German texts, the Italian translations were reviewed by individuals with intellectual disabilities to ensure their accessibility.

The source corpus comprises 26 German fact sheets (defined “Corpus Bolzano German”) and their Italian human translation (defined “Corpus Bolzano Italian”).

For the purpose of our study, we translated the “Corpus Bolzano German” into Italian with the help of DeepL, which led to the creation of the “Corpus DeepL Italian”. We used the free version of DeepL. The style was set to “automatic”. No post-editing was carried out. The “Corpus Bolzano German” (source corpus) contains a total of 12.416 words and 69.616 characters, while the “Corpus DeepL Italian” (target corpus) comprises a total of 15.453 words and 74.817 characters. The source German texts have an average length of 486,8 tokens, and the target Italian texts have an average length of 594,6 tokens.

The human translations of the “Corpus Bolzano German” (Corpus Bolzano Italian) served as gold standards for the evaluation of the DeepL performance. The evaluation followed the MQM criteria as adapted to Easy Language by Ahrens et al. (in press).

4 Results

Table 1 shows the quantification of the errors in the Corpus DeepL Italian compared to the gold standard texts (Corpus Bolzano Italian) following the MQM criteria. We followed the categorization put forward in Ahrens et al. (in press) with respect to the subcategories of the MQM core and added the category “incorrect explanation” (category: “accuracy”) for cases of incorrect or inappropriate explanations of technical or unfamiliar words. The category “Hallucination” was also added.

We used the category “Audience Appropriateness” for all issues related to deviations from the rules of Italian Easy Language. This is because such deviations make the text unsuitable for the target audience, either by increasing complexity and potentially causing misunderstandings or bearing the risk of stigmatization.

Errors and issues in the DeepL translation in Italian were annotated, evaluated and then discussed by both authors who have a native-level proficiency in Italian.

MQM Error type	Quantity in the Corpus DeepL Italian
Terminology	
Inconsistent terminology	15
Wrong term	52
Accuracy	
Mistranslations and semantic shifts	8
Hallucinations	2
Untranslated	24
Wrong explanation	7
Language conventions	
Grammar	7
Audience appropriateness <i>Deviations from the EL rules</i>	196
TOTAL	311

Table 1: The quantification of the errors in the Corpus DeepL Italian compared to the gold standard texts (Corpus Bolzano Italian)

5 Discussion

5.1 Terminology

Terminology-related issues regarding the DeepL translation were critical. The machine-translated text versions exhibited problems with the correct translation of technical or domain-specific terminology, where substituting synonyms would result in a loss of contextual clarity. This issue was particularly evident in the translation of names related to legal institutions, administrative bodies, services and professional titles. A specific difficulty arose in the translation of the names of administrative units in the municipality of Bolzano, which were generalized according to the German standard, causing the original terms in Italian to be omitted in the retranslation. For example, the “Sportello unico per l’assistenza e la cura” (“One-stop-shop for care and support”) was called “Punto di contatto per l’assistenza e il supporto” (“Contact point for care and support”) in the DeepL translation, while the “Sportello informativo per il cittadino” (“Citizen information point”) became the generic “Servizio al cittadino” (“Citizen service”).

The question of correct terminology also comprised abbreviations that were not translated in the target text but remained unaltered in their source text version, although they have a correspondence in the target text that is not identical to the source text. In a bilingual region like Bolzano, each language has its own set of abbreviations for the same institutions and processes, usually derived from their full forms. For example, the German abbreviation “E EVE” appeared in place of the Italian “DURP” in texts translated with DeepL. These untranslated abbreviations pose a risk of not being recognized or linkable to their full forms, especially if these full forms also appear in the text.

5.2 Accuracy

The evaluation of the Easy Language texts translated by DeepL revealed several significant issues related to accuracy. Semantic shifts due to incongruent synonymy were observed in the DeepL corpus. These errors arose when terms in the source language had a different scope or meaning, resulting in the use of inappropriate equivalents in the target language that did not align with the intended context. For example, this happened with the word “indennità” (“allowance”) becoming “paghetta” (“child’s pocket money”) and “amministratore di sostegno” (“legal guardian”) becoming “custode” (“guard”).

This also concerned the use of modal verbs in the translated texts. In several instances, these verbs were altered from their original form in the source text, resulting in substantial semantic shifts in both the German and Italian versions. An example of the Italian translation is the sentence from the Corpus Bolzano Italian “Anche le cooperative sociali devono guadagnare soldi” (Even social cooperatives must earn money) that appears as “Anche le cooperative sociali vogliono guadagnare soldi” (“Even social cooperatives want to earn money”) in the Corpus DeepL Italian.

Another category of errors involved non-translated sequences. In both translation directions, certain phrases remained unchanged from the source text, though this occurred in a very limited number of instances. For example, the names of Bolzano’s administrative units remained in German in the Italian text and in Italian in the German text, but not consistently. This inconsistency was also observed in the reverse translation direction, where different toponyms were either translated or left

untranslated compared to the other direction, showing a lack of systematic approach.

5.3 Hallucinations

Furthermore, the Corpus DeepL Italian displays some fragments of English with no relation to the source text. As the source text is in German they were labelled hallucinations. These ‘hallucinations’ in English create an additional obstacle to understanding complex content, such as that typical of administrative texts.

5.4 Untranslated fragments

In as many as 24 cases, untranslated fragments from the source text remained in the target text. They mainly concern toponyms for which both German and Italian terms are available. There is no consistency here in the target texts, which significantly reduces comprehensibility, especially for an audience with intellectual disabilities or other vulnerabilities in terms of understanding content.

5.5 Language conventions

Grammatical errors were identified in some of the machine-translated texts. While these errors were not numerous, they were recurrent within specific syntactic structures, affecting the overall grammatical accuracy of the texts.

5.6 Deviations from the rules of German and Italian Easy Language

DeepL’s translations of Easy Language texts do not adhere to the established German (Bredel & Maaß, 2016; Maaß, 2020) or Italian guidelines (Sciumbata, 2022; Perego, 2021), as the system is trained on standard and specialized language corpora rather than Easy Language rules. A qualitative analysis revealed several key deviations from these guidelines.

First, the translations tended to normalize the texts towards standard administrative language. This results in longer sentences and more complex vocabulary, often replacing simpler words with more high register synonyms, making the text less accessible. For example, the sentence from the Corpus Bolzano Italian, “Nel contratto di lavoro c’è scritto...” (“In the employment contract it says...”) is rendered in a more institutional tone in the Corpus DeepL Italian as “il contratto di lavoro stabilisce che...” (“The employment contract states that...”). Similarly, “La persona può lavorare”

(“The person can work”) becomes “la persona è idonea al lavoro” (“the person is suitable for work”) and “altre informazioni” (“other information”) appears as “ulteriori informazioni” (“further information”) in the DeepL translation. Second, the use of verbal tenses, modes, and voice does not align with Easy Language restrictions. While Easy Language guidelines limit German to the present and perfect tenses and discourage the subjunctive, and Italian similarly minimizes grammatical complexity, DeepL translations frequently include a broader range of tenses, including the conditional, gerund, future, and passive constructions. Here are two examples of translations that highlight these issues: “questa persona non può andare a lavorare” (“this person cannot go to work”) became “questa persona potrebbe non essere in grado di lavorare” (“this person might not be able to work”) and “gli esperti assistono le persone” (“experts assist people”) appeared as “le persone sono assistite da esperti” (“people are assisted by experts”). Another issue arose with impersonal constructions and double negatives. Easy Language favors action-oriented sentences that clarify actors and contact persons, avoiding impersonal and passive forms. Furthermore, negative statements and double negatives, which can obscure meaning, are discouraged. However, DeepL-generated texts frequently contained these structures, making the content harder to understand. For example: “lei trova le informazioni qui” (“you find the information here”) becomes “per informazioni si veda qui” (“for information see here” which is impersonal and grammatically requires the Italian subjunctive rendering this solution more complex in more respects) and “il libro è gratis” (“the book is free”) is “il libro non costa nulla” (“the book costs nothing”) in the DeepL translation. Consistency in terminology is also compromised. Easy Language guidelines require using the same term for the same concept throughout a document to enhance cohesion and clarity. DeepL translations, however, operate at the sentence level, failing to maintain consistency even within sections. For example in the same text, both “medico di famiglia” (“family doctor”) and “medico di base” (“general practitioner”) are used to refer to the same profession, whereas “parlamentari” (“parliamentarians”) are also called “membri del parlamento” (“members of parliament”).

Finally, the handling of difficult terms does not follow Easy Language principles. Technical or unfamiliar words should be explained when necessary, but DeepL does not adapt explanations to the difficulty level of terms in the target language. As a result, some complex terms remain unexplained when they should be, while others are unnecessarily explained, sometimes incorrectly, leading to a heavier and less effective text.

6 Conclusion and future work

The previous evaluation indicated that DeepL has achieved good, however not outstanding, results in interlingual translation of administrative texts into Easy Language. These texts present a risk of misinterpretation or misunderstanding among the target groups, primarily due to inaccuracies or a lack of compliance with Easy Language rules, which can hinder comprehension.

Moreover, administrative-legal communication is highly sensitive, requiring a level of accuracy comparable to that of health communication. As discussed in Deilen et al. (2023, 2024a, 2024b) for health-related texts, any content-related errors render a text unsafe. Consequently, automatic translation into Easy Language cannot be directly provided to the intended audience (e.g., via an institutional website) without human post-editing to ensure both accuracy and suitability.

This additional analysis using the MQM method will help identify where to prioritize post-editing efforts and facilitate comparisons (Lommel, Uszkoreit & Burchardt, 2014) with the results obtained from other artificial intelligence tools used for interlingual translation of Easy Language texts in the administrative domain.

References

- Sarah Ahrens, Silvana Deilen, Sergio Hernández Garrido, Ekaterina Lapshinova-Koltunski, Christiane Maaß. 2025 in press. Evaluation of Translations into Plain German Produced by Humans and MT Systems Including ChatGPT. *SKASE Journal of Translation and Interpretation*.
- Sarah Ahrens, Chiara Fioravanti. 2022. Cultural implications in Easy Language texts for migrants. Theoretical considerations and insights from practice in Germany and in Italy. *Trans-kom - Journal of Translation and Technical Communication Research*, 15 (2) 2022: 270–292.

- Ursula Bredel, Christiane Maaß. 2016. *Leichte Sprache. Theoretische Grundlagen, Orientierung für die Praxis*, Duden, Berlin.
- Silvana Deilen, Sergio Hernández Garrido, Ekaterina Lapshinova-Koltunski, Christiane Maaß. 2023. Using ChatGPT as a CAT tool in Easy Language translation. In *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria; 2023: 1–10. <https://aclanthology.org/2023.tsar-1.1>
- Silvana Deilen, Ekaterina Lapshinova-Koltunski, Sergio Hernández Garrido, Julian Hörner, Christiane Maaß, Vanessa Theel, Sophie Ziemer. 2024a. Evaluation of intralingual machine translation for health communication. In: *Proceedings of the 25th Annual Conference of the European Association for Machine Translation (Volume 1)*:469-479.
- Silvana Deilen, Ekaterina Lapshinova-Koltunski, Sergio Hernández Garrido, Christiane Maaß, Julian Hörner, Vanessa Theel, Sophie Ziemer. 2024b: Towards AI-supported Health Communication in Plain Language: Evaluating Intralingual Machine Translation of Medical Texts. In *Proceedings of the First Workshop on Patient-Oriented Language Processing (CL4Health)@ LREC-COLING 2024*, 44–53.
- Tira Nur Fitria. 2023. Performance of Google Translate, Microsoft Translator, and DeepL Translator: Error Analysis of Translation Result. *Al-Lisan: Jurnal Bahasa (e-Journal)*, Vol. 8.2, 2023.
- Walburga Fröhlich, Klaus Candussi. 2021. Easy Language in Austria. In Lindholm C, Vanhatalo U (eds.): *Handbook of Easy Languages in Europe* Frank & Timme, Berlin: 191-218.
- Abigail Kaplan 2021. *Suitability of Neural Machine Translation for Producing Linguistically Accessible Texts. Exploring the Effects of Pre-Editing on Easy-to-Read Administrative Documents*. Manuscript of the PhD thesis, University of Geneva.
- Camilla Lindholm, Ulla Vanhatalo (eds. 2021a). *Handbook of Easy Languages in Europe*. 2021. Frank & Timme, Berlin. DOI: 10.26530/20.500.12657/52628; <https://library.oapen.org/handle/20.500.12657/52628>
- Camilla Lindholm, Ulla Vanhatalo (2021b): Introduction. In Lindholm C, Vanhatalo U (eds. 2021a) *Handbook of Easy Languages in Europe*. Frank & Timme, Berlin: 11-26. DOI: 10.26530/20.500.12657/52628; <https://library.oapen.org/handle/20.500.12657/52628>
- Arle Lommel, Hans Uszkoreit, Aljoscha Burchardt. 2014. Multidimensional quality metrics (MQM): A framework for declaring and describing translation quality metrics. *Tradumatica*, 12:0455–463.
- Christiane Maaß. 2015. *Leichte Sprache*. Das Regelbuch, Berlin.
- Christiane Maaß. 2020. *Easy language – Plain Language – Easy Language Plus*. Balancing comprehensibility and acceptability, Frank&Timme, Berlin,
- Christiane Maaß. 2024. Intralingual Translation in Easy Language and in Plain Language. In Pillière L, Berk Albachten Ö; Hrsg (eds), *The Routledge Handbook of Intralingual Translation*, Routledge; London. doi:<https://doi.org/10.4324/9781003188872>
- Christiane Maaß, Chiara Fioravanti. 2025 in print, *Evaluating the performance of DeepL as translation tool between German and Italian Easy Language administrative texts*, RIID 1/2025.
- Christiane Maaß, Laura Marie Maaß. 2024, *Leichte Sprache bei intellektuellen Beeinträchtigungen*. In *Sprache - Stimme – Gehör, s.p.* DOI: 10.1055/a-2302-7802.
- Christiane Maaß, Isabel Rink. 2021. Translating legal texts into Easy language. In Chiara Fioravanti (ed) *Communicating the law and public information to vulnerable audiences*. JOAL Vol.9 n.1. <https://ojs.law.cornell.edu/index.php/joal/article/view/109>
- Christiane Maaß, Isabel Rink, Silvia Hansen-Schirra. 2021. *Easy Language in Germany*. in: C Lindholm, U. Vanhatalo (eds.), *Handbook of Easy Languages in Europe*, Frank & Timme; 2021. https://www.frank-timme.de/de/programm/produkt/handbook_of_easy_languages_in_europe?file=/site/assets/files/4477/2021_of_easy_languages_in_europe.pdf
- Christiane Maaß, Schwengber Laura Marie. 2022. Easy Language and Plain Language in Germany. In *Rivista internazionale di tecnica della traduzione International Journal of Translation*, 24: 43–61. doi: 10.13137/2421-6763/
- Anne Parpan-Blaser, Simone Girard-Groeber, Gabriela Antener, Christina, Baumann Rita, Alexandra Caplazi, Luisa Carrer, Cindy Diacquenod, Annette Lichtenauer, Andrea Sterchi. 2021. Easy Language in Switzerland. In Lindholm C, Vanhatalo U (eds.), *Handbook of Easy Languages in Europe*. 2021, Frank & Timme, Berlin: 573-622.
- Giulia Pedrini. 2024. *Medical communication between Plain Language and Einfache Sprache. A corpus analysis of layperson summaries of clinical trials in*

English, German, and Italian. Frank & Timme, Berlin.

Elisa Perego. 2021. Easy Language in Italy, In Lindholm C, Vanhatalo U (eds.), *Handbook of Easy Languages in Europe*. 2021, Frank & Timme, Berlin: 275 -304.

Isabel Rink. 2020. Rechtskommunikation und Barrierefreiheit. Zur Übersetzung juristischer Informations- und Interaktionstext. In *Leichte Sprache*, Frank & Timme, Berlin.

Carlotta Sciumbata. 2022. *Manuale dell'Italiano facile da leggere e da capire*, Franco Cesati editore, Firenze.