ChatGPT and Mistral as a tool for intralingual translation into Easy French

Julia Degenhardt
Université Bourgogne Europe / Maison de l'Université, Esp. Erasme, 21078 Dijon
Julia.Degenhardt@ube.fr

Abstract

FALC is a simplified variety of French designed to enhance text comprehensibility and accessibility. Despite its societal benefits, the availability of FALC texts remains limited due to the costly human translation process. This study explores the potential of LLMs, specifically ChatGPT and Le Chat, as a tool for automatic intralingual translations. The AI-generated translations of standard French texts on sexual health are compared to human-translated versions. The corpus-based study combines qualitative and quantitative approaches to evaluate content accuracy, readability and syntactic complexity.

1 Introduction

Like other Easy languages FALC (français facile à lire et à comprendre) is a complexity reduced variety of French, that follows guidelines to enhance text comprehensibility and accessibility (Lindholm & Vanhatalo, 2021). The French government promotes its use to improve societal inclusion of people with disabilities. In 2021 a charter on the accessibility of the communication between government and citizens was published, including recommendations to produce texts in FALC (Charte d'accessibilité, 2022). Despite, these efforts the number of texts available in FALC is rather small, the main reasons being high costs and difficulties to translate into FALC (Chehab et al., 2019).

Using generative AI to optimize the translation process could increase the text volume. Although scepticism and negative attitudes towards ChatGPT and other large language models do exist in the translation industry, the European Language Industry Study 2024 (ELIS) reveals that their use is growing. In 2024, 21% of the Language Service Companies had already implemented a LLM into their workflow (ELIS, 2024). In the study by Rivas Ginel and Moorkens (2024) 40% of the translators claimed they used ChatGPT regularly or occasionally, which further underlines the growing impact of LLMs in interlingual translation. This trend is backed up, by recent studies evaluating the proficiency of LLMs for the task. Although proficiency differs across models and languages, the results are promising and show potential (Jiao et al., 2023; Zhang et al., 2023; Zhu et al., 2024). LLMs have also been successfully tested on simplification tasks (Kew et al., 2023). Producing Easy Language texts is a form of intralingual translation and closely linked to text simplification. Yet, the potential of LLMs for this task remains largely unexplored. Deilen et al. (2023) explored the use of ChatGPT as a CAT tool for translations into Easy German. The authors obtained promising results, yet comparable studies for Easy French do not exist.

The present pilot study tries to address this research gap. The main question is how well ChatGPT and Mistral's LLM Le Chat can simplify a source text into FALC and whether they can be a useful tool especially for translators, but also for end-users. The analysis is a mix of qualitative and quantitative methods and focusses on content, readability and syntactic complexity. For this purpose, AI-generated translations – standard French into FALC – are compared to human-translated versions. The corpus contains 15 source texts in standard French on sexual health topics and their respective translation in three versions: human translator vs. ChatGPT vs. Le Chat.

The remainder of this paper is structured as follows: Section 2 reviews the current usage as well as the social and legal framework of FALC in France. Section 3 presents related work on Automatic Text Simplification in French and the

Proceedings of the 1st Workshop on Artificial Intelligence and Easy and Plain Language in Institutional Contexts (AI & EL/PL), pages 38–54, June 23, 2025

usage of LLMs for simplification and translation. Data and methodology are described in Section 4 and the results follow in section 5. To conclude, the main findings are summarised, and an outlook is given.

2 Easy Language

2.1 FALC – Easy French

Easy Language is an umbrella term for different simplified language varieties that aim at making information – mainly but not exclusively written texts – more accessible to different target groups with diverse communicative needs. The main target audiences addressed by Easy Languages are people with cognitive impairments or learning disabilities, aphasia, dementia, deaf or hard of hearing, functionally illiterate adults and foreign language learners (Lindholm & Vanhatalo, 2021b; Maaß, 2020). Easy Languages reduce the complexity on different linguistic levels (lexical, syntactical, discourse) in order to enhance comprehensibility and readability and to reduce the cognitive processing costs (Hansen-Schirra, Bisang, et al., 2020; Hansen-Schirra & Maaß, 2020). Producing Easy Language texts has been defined as a form of intralingual translation, which requires translational competences (Maaß, 2020).

Access to information is pivotal for inclusion and active participation in society, hence Easy Languages fall within the scope of accessible communication. Different texts can represent different communication barriers that hinder comprehension. That is for example the case if the text is not perceivable due to sensory impairments, if the language of the text in unknown to the reader or if the complexity of the content exceeds their cognitive processing capacities (Rink, 2019). Easy Language translation seeks to overcome these communication barriers in order to produce texts that are retrievable, perceptible, comprehensible, linkable, acceptable and action-enabling (Maaß, 2020). In France, Easy Language is known under the acronym FALC, which stands for "Français Facile à Lire et à Comprendre" and is commonly used in France, Belgium and Switzerland. Other terms referring to the same linguistic variety are "Français facile" or "Facile à Lire" (Canut et al., 2020; Vandeghinste et al., 2021).

2.2 Societal and legal framework

In 2005 the law (Loi n° 2005-102, 2005) on the rights of people with disabilities was passed by the French parliament. It is the most important legal text to date in France concerning equal rights, opportunities and participation. Article 47 of the law states that public authorities are obliged to make their online communication services accessible, however the text does not specify the means by which this aim is to be achieved. French Sign Language is the only form of accessible communication that is explicitly mentioned in the legal text (Loi n° 2005-102, 2005). Consequently, there is no legal framework regarding texts in FALC in France, as it is the case in Germany. This might be one of the reasons why providing simplified versions is still rather an exception than the rule.

Nonetheless, awareness of accessible communication has grown in recent years. In 2021 the French government published the first version of the "Charte d'accessibilité de la communication de l'État", which recommends providing additional texts in FALC (Charte d'accessibilité, 2022). The charter specifically mentions electoral programs as one of the document types that should be made available in FALC. This was implemented by a great number of candidates during the election campaign in 2022. Furthermore, a campaign with the headline "*Imaginer un quotidien où rien n'est vraiment pensé pour vous*" (engl.: "Imagine a daily life where nothing is designed for you.") was launched in 2025. The aim is to raise awareness on accessibility amongst public agents in general, but also to enhance knowledge about specific measures for communicative inclusion like FALC (Ministère du Travail, de la Santé, des solidarités et des familles, 2025). Most of the texts currently available in FALC are informative, focusing on areas such as disability, inclusion, healthcare, political participation and cultural events, for example in the form exhibition guides for museums (Chehab et al., 2019).

2.3 Text production

Research activities on Easy Language varieties in France have also lagged behind those in other European countries, including Germany, Spain, Finland, where research activity but also the number of available texts has been increasing (for an overview see Lindholm & Vanhatalo, 2021a). Although the French guidelines for producing texts

in FALC have available since 2009, a survey on its use amongst public and private organizations in 2019 revealed that producing FALC texts is seen as time-consuming and too difficult. The organisations surveyed are aware of the necessity, but they often do not know how to integrate FALC translation processes into their workflow (Chehab et al., 2019). In France the professionalisation of the field is in its early stages, professional translators are rare, but demand is growing. This situation suggests that there is a growing need to optimize the translation process, incorporating at least some degree of automation.

3 Related Work

3.1 Automatic Text Simplification in French

Text simplification can be generally defined "as the process of reducing the linguistic complexity of a text, while still retaining the original information content and meaning." (Siddharthan, 2014, p. 259). Automatic Text Simplification (ATS) has been researched for years, not only to produce readable texts for humans but also as a form of pre-processing for other NLP tasks. As there is a great need for simplified texts in order to enhance societal inclusion, provision has become increasingly important (Saggion, 2024).

Most of the early research on ATS was carried out in English and corpus data on other languages like French was scarce, which hindered the development of performant statistical, rather than less performant rule-based, tools for simplification in French. Seretan analyzed the simplification strategies adopted by human translators and derived a ruleset for syntactical simplification in French from the results (Seretan, 2012). Brouwers et al. described the main linguistic levels of transformation: lexical, discursive and syntactical and incorporated them into a rule-based system. This approach obtained good results, with about 80% of the generated sentences being correct (Brouwers et al., 2014).

In recent years, interest has shifted towards machine-learning approaches and much research has been dedicated to the construction of French parallel corpora to address the lack of data. Ormaechea & Tsourakis created the open-source Wikipedia Vividia Corpus (WIVICO 10) by extracting and aligning complex/simple sentence pairs from comparable corpora (Ormaechea & Tsourakis, 2023). They also addressed the problem that simplified sentences can still exhibit complex structures and that complexity evaluation does not always account for this. Most evaluation measures can only identify whether the generated sentence is simpler, but not to which degree. As 'simpler' does not immediately equal maximum simplicity and comprehension, this is problematic for the evaluation of ATS tools. To improve the assessment of sentence complexity, the authors fine-tuned a pre-trained BERT classification model. Results showed that their model is useful for automatic creation of simplified datasets as it provides a finer-grained assessment of simplification (Ormaechea & Tsourakis, 2024). Another available French corpus that has been used to evaluate ATS systems, is the ALECTOR corpus created by Gala et al. (2020). It contains literary and scientific texts conceived for elementary school children and their respective simplified versions. Simplified versions were created manually by applying simplification strategies on lexical, morphological and syntactical level. Although initially collected to assess reading errors and to improve reading skills in young children with dyslexia, it is also useful for ATS (Gala et al., 2020). ALECTOR served as the basis to develop the French ATS system HECTOR. This system combines a rule-based and an embedding-based approach to perform simplification at lexical, syntactical and discursive level. Given the focus of the corpus data, it has a strong focus on learner texts for young children. The researchers obtained good results for syntactical simplification, but the system was less powerful at lexical and discursive level (Todirascu et al., 2022). The CLEAR corpus, which comprises original and simplified texts in French from the medical domain, has also been used to address automatic sentence extraction and alignment (Cardon & Grabar, 2019; Grabar & Cardon, 2018). This small specialized corpus also provided data for a later study by Cardon & Grabar, where they showed that that high quality specialized data and translated corpora can be successfully used to train ATS models, even if performance will increase in line with the size of the data set (Cardon & Grabar, 2020). These findings were confirmed by Abdul Rauf et al. (2020), who used a synthetic corpus, consisting of the French translations of English source texts of the Newsela corpus, to train their simplification model. Although their results varied across the different levels of complexity, the authors' overall conclusion was that small data batches and

translated corpora can result in acceptable simplifications (Abdul Rauf et al., 2020). While the previous mentioned ATS models explored simplification on various text levels, the *FrenLys* tool investigates lexical simplification. It generates, selects and ranks synonyms to replace complex words in a text. (Rolin et al., 2021).

3.2 LLMs for intralingual translation tasks and simplification

Easy Language translation is a form of intralingual translation. While research on the former is scarce, many studies have assessed the capabilities of LLMs for interlingual translation tasks. The results are heterogeneous but promising, showing that performance depends significantly on the model, the languages and the prompts used. Especially for high-resource languages, LLMs can produce qualitatively good and competitive outputs (Hendy et al., 2023; Jiao et al., 2023; Zhu et al., 2024). According to Vilar et al., who tested the MT capabilities of an LLM against state-of-the art MT systems, the LLM "matches the fluency but lags the accuracy of conventional NMT" (Vilar et al., 2022). Despite some weaknesses, the usefulness of LLM interlingual translation has been demonstrated, suggesting that such approaches may also produce useful results for intralingual tasks.

Besides interlingual MT, the simplification capacities of LLMs have also been assessed. Feng et al. performed sentence simplification using ChatGPT amongst others and concluded that "LLMs outperformed current state-of-the-art [sentence simplification] methods." (Feng et al., 2023). In regard to text simplification, Kew et al. also concluded that LLMs perform better than state-of-the art text simplification baseline models (2023). Furthermore, these findings are confirmed by Qiang et al. who claim that the GPT-4o model "not only simplifies text effectively but also produces output that is easier to read." (Qiang et al., 2025). Although text simplification and intralingual translation into Easy Languages are not the same (different target groups, specific rule set, etc.), reducing complexity is crucial for both operations. Thus, one can hypothesise that LLMs do not only perform well in ATS but also in Easy Language translation. Yet, their potential remains mostly unexplored. For Easy German, Anschütz et al. (2023) and Klöser et al. (2024) demonstrated that pre-training LLMs with Easy Language data combined with fine-tuning results in models that can produce satisfying Easy German texts. Deilen et al. (2023) examined the usability of ChatGPT as a CAT tool for intralingual translation of administrative texts into Easy German. The author's results were promising: ChatGPT produced texts that were simpler on some linguistic levels but also contained content errors. Hence, they concluded that ChatGPT can be useful but not without post-editing (Deilen et al., 2023). Arguably, using LLMs or other ATS tools for Easy Language text production is of great interest, because it might save time and money, two factors which are often named as major impediments for Easy Language translations (Chehab et al., 2019). Increasing the number of texts produced in Easy Language plays a crucial role in the efforts to make society more accessible. The social dimension of Easy Language translation is also a driver of research on the automatization of the process (Saggion, 2024). Although it comes with its challenges, the use of machine translation, terminology management, etc. has become increasingly important for intralingual as well as interlingual translators (Hansen-Schirra et al., 2020). LLMs hold a large potential as they are free and easy to use. However, for French this potential remains currently unexplored. This pilot study is a first approach to bridge this research gap and to initiate a discussion on using LLMs as a tool for producing texts in Easy French.

4 Methodology

4.1 Data Collection

The present study is based on a French monolingual corpus. It consists of original source texts (ST) in standard French and the translated target texts (TT) of these STs in FALC in three different versions. The different versions of these TTs are:

1. official TTs translated by human translators, that were published on the websites alongside the standard French STs. These texts were collected as part of the corpus.

2. TTs that were generated by the author using two different Large Language Models.

The LLMs chosen for this study are ChatGPT (version 4o mini) by OpenAI and Le Chat (version

Mistral Large) by Mistral AI. ChatGPT seems like an obvious choice due to its popularity, the user-friendly interface and free subscription. Furthermore, other studies in the field have already discussed ChatGPT's potential for intralingual (Deilen et al., 2023) and interlingual (Jiao et al., 2023) translation, and prompting strategies have also often been tested on ChatGPT (Campesato, 2024; Gao et al., 2024). Le Chat is very similar to ChatGPT: both are free, and the user interfaces hardly differ from each other as they are dialogue-based. Although it is certainly less popular on an international scale than other LLMs like Google's Gemini, Mistral AI is one of the most successful European AI companies. The French-based company signed a contract with Microsoft in 2024, which further increased its market value (Braune, 2024). Since public agents are amongst the groups for whom using an LLM for translations into FALC might be beneficial, the fact that France Travail (the French public employment service) already is one of Mistral AI's clients was another argument for choosing Le Chat (Mistral AI, n.d.).

The STs are informative texts from the medical domain [1]. Most texts concern sexual and reproductive health subjects and are targeted at young adults, while some texts aim to inform a broader audience about mental health or breast cancer. While the source texts include domain-specific language, they are written for lay people and not domain experts. All texts were originally published in France between 2019 and 2024 and are freely available online.

The main selection criterion for the texts was that a clear link between the target text in FALC and the source text in standard French could be established. As mentioned above, this is rarely the case in France – most FALC texts available online are not labelled as translations and cannot be traced back to a source text (Chehab et al., 2019). In that respect, it is also difficult to get information about the professional background of the translators. It is more likely that they are working in the disability field than as professional translators (ibid.). Some of the texts have been produced in cooperation with associations for people with disabilities. However, it remains unclear whether their role relates to consultation, translation or proofreading. Ideally, this information would be included within the corpus metadata, but it is not available. Furthermore, the target texts had to be comparable in terms of domain and subject. Thus, texts about other subjects than health were excluded from this study. Those criteria clearly limit the number of eligible texts. Considering that the number of texts in FALC is already small, some compromises in the collection process were necessary to increase the sample size (Chehab et al., 2019; Rodríguez Vázquez et al., 2022). On the one hand, this concerns the text length, which differs. On the other hand, this concerns the lack of metadata, especially regarding the professional background of the translators. However, restricting the selection to texts of similar length or to the availability of metadata would not have yielded a sufficiently large corpus.

To summarize, the corpus consists of 32214 words in total, distributed across four subcorpora. Each subcorpus contains 15 texts. ST_StFR contains the STs in Standard French. The TTs in FALC are categorized according to the translation process: human translators (TT-1_human) vs. LLM-generated versions (TT-2_ChatGPT and TT-3_LeChat). Table 1 shows the number of words in each subcorpus.

subcorpus	words
ST_StFR	10143
TT-1_human	9705
TT-2_ChatGPT	7490
TT-3_LeChat	4876
total	32214

Table 1: Corpus Statistics

4.2 Prompting Strategies

LLMs generate their output based on the prompt provided by the user. The quality and structure of the prompts plays a crucial role and affects the output. Different prompts will produce different responses, and the same prompt will not reproduce the same answer. The more precise and well-structured the prompt the more concise the output will be. Especially for complex tasks, well-designed prompts are pivotal. In general, instructional and guided prompts that give clear instructions and provide additional context produce

[1]Please see the appendix for a list of the source texts and the respective links.

more precise output than open-ended prompts (Campesato, 2024).

This holds also true for translation tasks. Here context helps the model to better resolve ambiguity and choose suitable equivalents based on the provided context (Campesato, 2024; Hui Jiao et al., 2024). The benefits of assigning a role to the model are well-known and again, clarity is key. For translation tasks, assigning the role of a translator instead of just an author yields better results (He, 2024). Other studies have shown that providing domain specific information, such as indicating the translation direction, the style and text type of the translated texts, the text function and the target audience, tends to improve the quality of the target texts (Gao et al., 2024; Hui Jiao et al., 2024; Yamada, 2023). All these findings were considered for the prompts used in this study. The initial prompt[2] includes the following key information:

- role: translator
- task: simplify according to the FALC rules; the basic principles, e.g. short sentences, active voice, explication of complex words, were introduced in the prompt to provide context to the task
- direction: intralingual, standard French to Français Facile à Lire et à Comprendre (FALC)
- target audience: people with reading difficulties
- domain & text type: informative, sexual health

In their study on ChatGPT as a CAT tool for Easy German, Deilen et al. (2023) compared two different prompts. One approach was to break down the simplification process into linguistic levels. Although this prompting strategy complies with the finding that step by step-instructions are beneficial (Hui Jiao et al., 2024), this technique was not adopted here, because it is more time-consuming and it did not outperform the holistic approach in each category (Deilen et al., 2023). As iterations are recommended (Campesato, 2024), ChatGPT and Le Chat were asked three times to simplify the text. The second and third prompt asked the models to further simplify the text they just produced by keeping the rules of FALC in mind. The third simplified version was integrated in the corpus and analysed.

4.3 Data analysis

4.4 Content

A qualitative analysis of five source texts and their respective target texts was done manually. The chosen texts are about abortion, menstruation, sexually transmitted diseases, contraceptives and breast cancer screening. The analysis focusses on information consistency, added explications and content errors. The concept of a faithful delivery of the original message and information consistency are often seen as ideals in the context of automatic text simplification (Siddharthan, 2014) and Easy Language translation (Maaß & Rink, 2020). However, there is a risk of informational overload for the target audiences of Easy Language when the text contains too much information and becomes too long. The translators need to cut out non-essential information in order not to exceed the cognitive processing capacities of the readers (Maaß & Rink, 2020). Consequently, omissions cannot be counted as content errors in general. Easy Language translation settings are often characterised by an asymmetry in knowledge and translators face the challenge of bridging this gap and building common ground between producer and reader. Thus, adding information is as necessary as reducing information. The challenge is to decide whether information is crucial or not. For that matter, knowledge about the target group is a necessary competence for the translators to make adequate decisions (Hansen-Schirra, Bisang, et al., 2020; Maaß, 2020). The question is, then, whether LLMs are also capable of making these choices or whether too much information is omitted. The resulting hypothesis is that LLMs omit more information than the human translator and that they produce more errors, due to hallucinations, as it was the case with ChatGPT for Easy German (Deilen et al. 2023).

4.5 Readability

The readability was assessed through different measurements. First, the Moving-Average Type-Token-Ratio (MATTR) was calculated for each text. A lower MATTR indicates less lexical diversity and consequently higher readability. In

[2] Please see the appendix for the entire prompt

contrast to the TTR, which highly depends on text length, the MATTR is insensitive to text lengths as it calculates the type-token ratio over a sliding window (Covington & McFall, 2010). It has been demonstrated that MATTR is a reliable index to measure lexical diversity (Bestgen, 2024; Kettunen, 2014). The window-size was set to 50 tokens[3]. Secondly, the lexical density (LD) was computed. It describes the proportion between content and grammatical words in a text. A lower LD is an indicator for higher readability (Baker, 1995). Lastly, the AMesure-score was used to assess the overall readability. AMesure is a readability measurement tool for French language, initially designed to assess administrative texts. It takes into account various parameters of readability (e.g. lexical density, type-token-ratio, sentence length, verbal forms) to evaluate a text on a scale from 1 to 5 – the lower the score the more readable (François et al., 2014; François et al., 2020).

4.6 Syntactical complexity

Syntactic simplicity contributes to the comprehensibility of a text (Christmann & Groeben, 2019). The FALC guidelines recommend short sentences that only express one idea. Subordinate clauses should be avoided (Inclusion Europe, 2009). A smaller amount of dependency relations indicates lower complexity (Deilen et al., 2023; Deilen et al., 2024). Consequently, the TTs are expected to contain fewer complex clauses than the STs.

To evaluate the syntactical complexity of the target texts, the dependency parser from the Stanza NLP Library was used (Qi et al., 2020). Stanza extracts dependency relations as described in the Universal Dependencies (UC) framework (Marneffe et al., 2021). Based on Deilen et al., 2023 the following dependency relations were selected for the analysis: acl (clausal modifier of noun), acl:recl (relative clause modifier), advcl (adverbial clause modifier), aux:pass (passive auxiliary), appos (appositional modifier), ccomp (clausal complement), xcomp (open clausal complement), nsubj:pass (passive nominal subject), parataxis.

5 Results

5.1 Content

In the small selection of 15 target texts (human translator, ChatGPT, LeChat) no content error was detected. This finding is not consistent with the results by Deilen et al., who found at least one piece of incorrect information in over 60% of the ChatGPT texts (2023).

The qualitative content analysis did not confirm the hypothesis: human translators were not more consistent than the LLMs; on the contrary, the LLMs omitted less information units, as table 2 shows.

	Total counts of information units			
Text	ST	TT-1	TT-2	TT-3
Menstruation	50	41	46	39
Abortion	44	43	39	24
Contra-ceptive	56	42	50	45
IST	83	48	71	65
Breast Cancer screening	119	36	83	56
	352	210	289	229

Table 2: Number of information units

The most striking discrepancy concerns the brochure on breast cancer screening. If the information units in the ST are compared to those included in the human translation, 2/3 were omitted. These omissions are for example: symptoms for breast cancer are not explained, none of the statistics mentioned in the ST were cited in the TT, difference between benign cysts and cancers is not explained. Despite the fact that some of those information units could be classified as crucial, the TT does include much information about the screening procedure, which is not included in the ST. The focus of the texts shifted. While the ST is more general and gives some information about early symptoms and why and how to do a screening, the human TT is very specific about the screening but completely omits the symptoms. Such a shift in focus was only detected in this case, all the other analysed TTs kept the main subject.

[3] Covington & McFall, 2010 do not recommend a specific window-size, but Bestgen, 2024 found that 50 is common.

As the numbers in Table 2 show, the simplified versions include less information than the ST. This is in line with the FALC requirements: omissions are necessary to not overstrain the processing capacities of the target audiences (Hansen-Schirra et al., 2020). The following examples [4] will illustrate some cases of omissions.

Example 1: The source text explains early signs of a pregnancy.

1. Le premier indicateur d'une grossesse est souvent un retard de règles. Tu peux aussi avoir d'autres signes : nausées, mal à la poitrine, ventre gonflé… [The first indicator of pregnancy is often a late period. You may also have other signs: nausea, chest pain, a swollen belly,…] – ST-StFR

Le Chat translated this part as follows:

2. Comment savoir si on est enceinte ? Faites un test de grossesse. [How to know if you are pregnant? Take a pregnancy test.] – TT-3_LeChat

Nothing is said about early symptoms, which is a complete omission. This kind of information loss is problematic, because the reader is not well informed. It also negatively affects the coherence of the text, as the link between cause (early pregnancy signs) and consequence (take a test) is not clearly established as it is the case in the ST. ChatGPT and the human translator on the other hand translate the cause-consequence relation consistently as:

3. Un retard des règles peut être un signe de grossesse. Tu peux aussi avoir : des nausées (mal au ventre), des douleurs dans la poitrine, un ventre gonflé. [A late period can be a sign for pregnancy. You may also have: nausea (belly ache), pain in your chest, a swollen belly.] – TT-2_ChatGPT

4. Pour savoir si tu es enceinte, il y a plusieurs signes: tes règles sont en retard, tu as la nausée, tu as mal à la poitrine, tu as le ventre gonflé… [There are several signs that you may be pregnant: your period is late, you feel nauseous, your chest hurts, your stomach is swollen...] – TT-1_human

Example 2: The ST on menstruation states the following:

1. Si tu as d'autres symptômes douloureux qui t'empêchent de faire tes activités habituelles (douleur jusqu'à vomir, évanouissements...), il se peut que tu souffres d'endométriose. N'hésite pas a consulter. [If you have other painful symptoms that prevent you from doing your usual activities (pain to the point of vomiting, fainting, etc.), you may be suffering from endometriosis. Don't hesitate to get a consultation.] – ST_StFR

The Le Chat (2) and the human TT (3) are both less specific, Le Chat does not even mention endometriosis. Only ChatGPT (4) omits no information:

2. Si tu as beaucoup de douleurs, parle à un médecin. [If you have a lot of pain, speak to a doctor.] – TT-3_LeChat

3. Si tu as vraiment très mal, tu peux aller voir un médecin. Tu as peut-être une maladie, qu'on appelle l'endométriose. [If you're in really bad pain, you can go and see a doctor. You may have a condition called endometriosis.] – TT-1_human

4. Si la douleur est très forte (par exemple, vomir ou s'évanouir), cela peut être un signe d'endométriose. Cela signifie qu'il faut consulter un médecin. [If the pain is severe (e.g. vomiting or fainting), this may be a sign of endometriosis. This means that a doctor should be consulted.] – TT-2_ChatGPT

Example 3: The source text explains that dropping hormone levels are what causes the body to evacuate the uterine lining at the end of each menstrual cycle if no egg is fertilized. While each target text explains that the body expels the uterine lining when fertilization has not occurred, none of

[4] Examples are originals taken from the corpus. However, the original layout of the FALC texts (one line, one sentence) was not maintained here.

them mentions that falling hormone levels are the cause.

1. Si l'ovule n'est pas fécondé, l'utérus se vide. [If the egg is not fertilized, the uterus empties.] – TT-2_ChatGPT

2. Si l'ovule n'est pas fécondé, l'utérus se débarasse de sa muqueuse. [If the egg is not fertilized, the uterus sheds its lining.] – TT-1_human

Example 4: The ST about sexually transmitted diseases explains that HP-viruses can be benign but some types might cause cancer. The LLM generated TTs do inform about the cancer risk, but not about benign forms. The human translator omits both information units.

1. Certains HPV peuvent causer des cancers. Un vaccin existe pour les éviter. [Some HPVs can cause cancers. A vaccine exists to prevent them.] – TT-3_LeChat

2. Les papillomavirus : Il existe un vaccin. [HPV: a vaccine exists.] – TT-1_human

Example 5: The ST about the morning-after pill explains the time frame for effective use, but the human translator omitted that information unit completely, in both LLM versions it is included:

1. Il faut prendre la contraception d'urgence. Tu peux la prendre jusqu'à 5 jours après le rapport. [You need to take emergency contraception. You can take it up to 5 days after intercourse.] – TT-2_ChatGPT

2. Prenez la pilule d'urgence dès que possible. Vous avez jusqu'à 5 jours pour la prendre. [Take the emergency pill as soon as possible. You have up to 5 days to take it.] – TT-3_LeChat

3. Il faut la prendre le plus tôt possible après un rapport à risque. [Take it as soon as possible after unprotected intercourse.] – TT-1_human

Example 6: The ST on sexually transmitted diseases explains three different types of screening, e.g. blood analysis. However, the human translator only lists two of the methods, while ChatGPT and Le Chat included all three.

1. Selon l'IST, le test peut être différent (sang, urine, auto-prélèvement). [Depending on the STI, the test may be different (blood, urine, self-sampling).] -TT-3_Le Chat

2. Tu peux aussi aller voir ton médecin, puis aller dans un laboratoire, où on testera ton sang, ou ton urine. [You can also see your doctor, then go to a laboratory, where your blood or urine will be tested.] – TT-1_human

These examples illustrate cases of complete omission. On the one hand, some can be rated as adequate omissions, e.g. example 3 and 6, on the other hand, in examples 4 and 5 crucial information is missing. Omissions always entail information loss, but these examples show that it is a gradable phenomenon. Reducing the amount of information is a common and necessary translation strategy (Hansen-Schirra et al., 2020; Maaß & Rink, 2020). Yet the decision often implies some degree of subjectivity, and the qualitative analysis shows that it is a problem for the translators and the LLMs.

Regarding the explanation of difficult concepts or words, the results are mixed. ChatGPT tends to add small explanations in brackets after a difficult word. While it is positive that the difficulty of a word was acknowledged, the format does not comply with the rules for FALC. More substantial explanations can be found in the texts translated by the human translator. For instance, in the text about the menstrual cycle, a whole paragraph was added, explaining what the period is: "Quand tu es une femme, ou une personne avec un utérus, tu peux avoir tes règles. L'utérus est un organe du corps humain. Quand tu as tes règles, du sang coule à l'extérieur de ton vagin. C'est naturel. Les règles font partie d'un cycle du corps, qu'on appelle le cycle menstruel." [When you are a woman, or a person with a uterus, you can have your periods. The uterus is an organ in the human body. When you have your period, blood flows out of your vagina. This is natural. Menstruation is part of a cycle in the body called the menstrual cycle.] – TT-1_human. Those kind of long explanations and additions have not been found in the TTs generated by the LLMs, although the prompt specified to add explanations if necessary.

5.2 Readability

Table 3 shows the Moving-Average Type-Token-Ratio and the lexical density for each subcorpus. As expected, the standard French STs have a higher mean MATTR than the TTs, indicating that the vocabulary used in the FALC texts is less diverse and, consequently, the texts are less complex. Amongst the TTs, the human versions have the lowest mean MATTR with 0.689 and ChatGPT produced the texts with the highest value.

These mean lexical density scores are interesting. One would expect a decrease from the STs to the TTs, but this only the case for the human translated TTs. Le Chat produced TTs that are denser than the STs and hence, presumably more complex.

Moving-Average Type-Token-Ratio (MATTR)			
subcorpus	mean value	highest value	lowest value
ST_StFR	0.757	0.797	0.69
TT-1_human	0.689	0.74	0.64
TT-2_ChatGPT	0.735	0.76	0.7
TT-3_Le Chat	0.702	0.741	0.615
Lexical Density			
ST_StFR	52%		
TT-1_human	48%		
TT-2_ChatGPT	52%		
TT-3_Le Chat	55%		

Table 3: MATTR and Lexical Density

The AMesure score was not as informative as expected, as all the STs scored 2 out of 5 (1 corresponds to the lowest complexity level), except for one text with a 3, indicating that the source texts already had a low level of complexity. The majority of the TT versions obtained the same score as the STs. All three TT versions of the text on violence in relationships, categorized as level 3, improved by one level. Most of the other TTs obtained the same score as the STs. This does not mean that the target texts have not been simplified at all, but rather that they have not been simplified sufficiently to change the overall score. As the AMesure score measures different parameters and weights them according to their impact on text complexity, it is probable that the simplifications made did not have enough weight to change the score (François et al., 2020).

5.3 Syntactical complexity

The analysis of the syntactic complexity shows that the source texts have the highest number of words per sentence with an average of 16. Le Chat produces the shortest sentences, with only 7 words/sentence on average. The source texts also have the smallest number of sentences in total, which is not surprising, as one important rule in FALC is to write short sentences and to split complex hypotactic sentences. As table 4 shows ChatGPT and LeChat are roughly similar in terms of total number of sentences, but not regarding the average sentence length. The corpus in the pilot study is too small to generalize but it seems that LLMs tend to produce shorter texts than human translators.

subcorpus	sentences in total	words/sentence
ST_StFR	633	16.02
TT-1_human	919	10.56
TT-2_ChatGPT	676	11.08
TT-3_Le Chat	667	7.31

Table 4: Sentence length

When comparing the relative frequencies of all examined dependency relations combined, complex clausal relations are most frequent in the STs. The TTs by ChatGPT, the human translators and Le Chat follow in descending order. Overall, the TTs contain less of the examined dependency relations, as Figure 1 illustrates. According to Deilen et al. (2023) decreasing frequencies of complex clausal relations indicate that the text is easier to understand.

The distribution of the different dependency relations over the subcorpora varies a lot. Even though the STs have higher counts in total, they do not exceed the TTs in every category. For instance, the human TTs include more clausal complements (ccomp) and more adverbial clause modifiers (advcl) than the STs. As subordinate clauses should be avoided according to the FALC rules, it is surprising that some of the clausal structures analysed are even more frequent in the TTs than in standard French. Open clausal complements (xcomp) are the most frequent dependency relation in the STs, the human TTs and the ChatGPT TTs. Xcomp-relations are core arguments of the verb, but without their own subject: as such, they often appear when modalities are expressed. Since the modal verb “pouvoir” (can) is either the second or

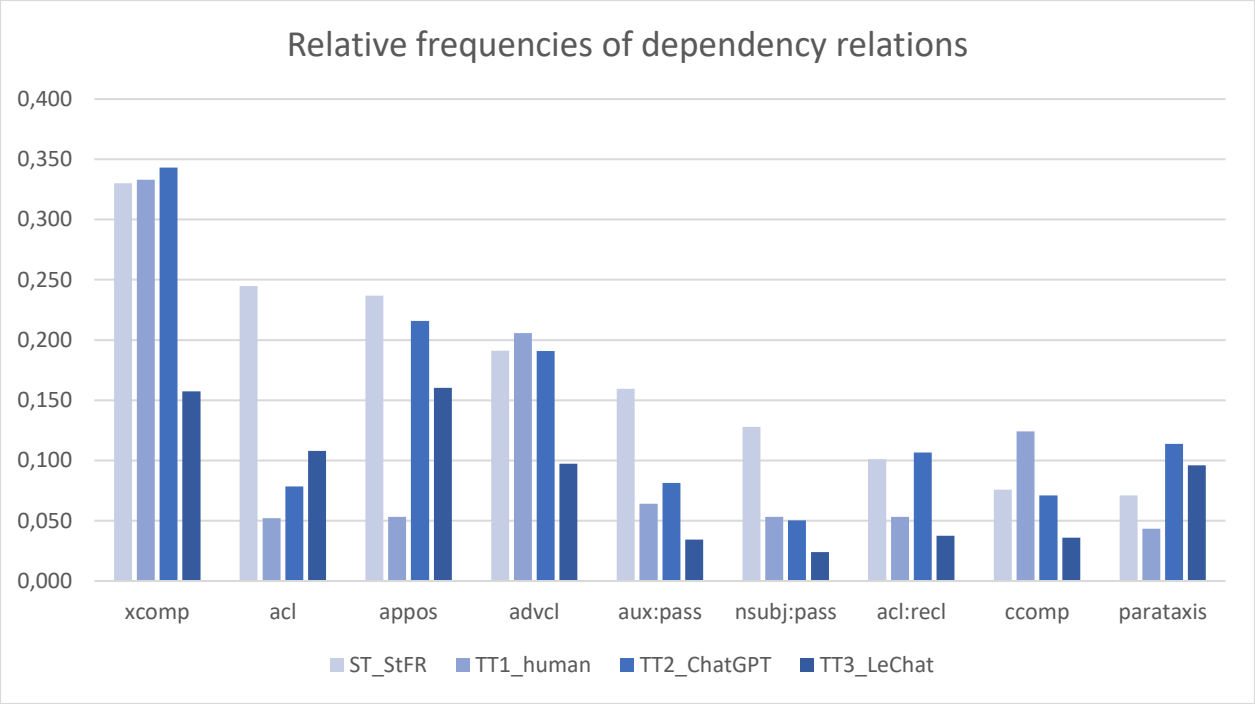

Figure 1: Relative frequencies of dependency relations

third most frequent verb in the subcorpora, the high number of xcomp-relations is not surprising. The following examples from the TT-2_ChatGPT subcorpus illustrate this:

1. Le cancer peut prendre de temps pour se développer. [Cancer can take time to develop.]

2. Cela peut durer plusieurs mois ou années. [This can take several months or years.]

3. Tu peux dire non.[You can say no.]

4. Si une femme enceinte ne veut pas garder son bébé, […].

5. [If a pregnant woman does not wish to keep the baby, […]]

Adverbial clause modifiers (advcl) have nearly the same relative frequency in the STs and in the ChatGPT TTs and the number is slightly higher in the human TTs. The similarity of these numbers is unexpected, as subordinal relations are not permitted by the FALC guidelines. The following examples from the TT-2_ChatGPT subcorpus illustrate the use of these clauses:

1. *Tu as le droit de dire non, si tu ne veux pas.* [You have the right to say no, if you don't want to.]

2. *Le cancer peut prendre de temps pour se développer. [Cancer can take time to develop.]*

3. *Trouver le cancer tôt permet de mieux le soigner. [Finding cancer early means better treatment.]*

4. *Il est important de commencer rapidement, pour respecter les délais. [It's important to get started quickly, to respect the deadlines.]*

The human-translated target texts include more clausal complements (ccomp) than the LLM-versions and the STs. One explanation for these higher numbers is that formulations such as "ça veut dire", "ça signifie" are used frequently to explain difficult words or concepts. Here are some examples (from TT-1_human):

1. Cela veut dire qu'ils sont secrets. [That means they are secret.]

2. Vous allez voir votre médecin cela s'appelle une consultation. [You will see your doctor, that is called a consultation.]

Clausal complements are also part of the construction "il faut X". The frequency per million tokens of the verb "falloir" is 3040 in the human TTs against 1958 (TT-2_ChatGPT), 899

(ST_StFR) and 640 (TT-3_LeChat). This explains why ccomp-relations are more frequent in the TT-1_human subcorpus.

6 Conclusion and future directions

The initial research question was whether ChatGPT and Le Chat could translate a source text into FALC and whether the output could compete with a target text that was translated by a human. Human translations are still seen as the gold standard for Easy Language translation. This is not only because automatic simplification tools either do not exist for a specific language or do not produce the desired outcome, but especially because they lack the ability to account for the different communicative needs of the very heterogenous target audience of Easy Language (Saggion, 2024). The necessary competences for an Easy Language translator include knowledge of the target audience to be able to adapt the content – both by adding and reducing the information appropriately (Maaß, 2020). We might expect that human translators are more capable of judging which information to include. However, the qualitative analysis did not confirm this, the LLMs were in some cases more consistent and omitted less information, while the human translators sometimes omitted relevant information. For example, the human translator omitted information about the time span for taking the morning-after pill, while ChatGPT and Le Chat did not. Although this is just one example, it demonstrates that assessing the adequacy of omissions is not only very difficult, but also that human judgement is error-prone. Therefore, potential content inconsistencies between ST and TT are not a sound basis to judge the capacity of ChatGPT or Le Chat to translate into FALC. As the qualitative analysis showed, the LLMs did not produce incorrect information and most of the information units was translated. Now, if we assume that a standard French ST gets translated by an LLM into FALC, we can look at the product from two perspectives: that of end user- and translator. The motivation to translate the texts differs: the user seeks information and needs a simplified version of the ST; the translator might seek inspiration or want to save time. From a user perspective, if crucial information is missing, the text might not be action-enabling as it should be (Maaß, 2020). Easy Language target audiences are unlikely to be able to search for the missing information elsewhere. Although the text might fail to enable its reader to act, based on the findings in this study, it is likely that the LLM produces a simpler text (in terms of readability and syntax), which can be interpreted as an improvement over the inaccessible ST. The situation is obviously different for translators, because they are not the end-users. If information units are missing, the translator can add them.

The results presented show that the question of whether LLMs are useful tools for FALC cannot simply be answered with yes or no. Yes, because overall the LLMs produced simpler versions of a source text. The sentences were shorter, the MATTR and the lexical density was lower (except for Le Chat) and the overall syntactic complexity decreased. Also yes, because the overall content was consistent despite some omissions. On the other hand, some of the dependency relations are more frequent in the target texts than in the source texts. This is for instance the case for adverbial clauses and open clause complements. The question is, then, to which extent each individual type of dependency relations affects the overall syntactic complexity for the target groups. Yet, this is a research desideratum, that has not yet been answered (Hansen-Schirra et al., 2020). In her study on the comprehensibility of clausal sentences in Easy German, Borghardt found that splitting them into two sentences to avoid subordination does not enhance the comprehensibility and, moreover, conjunctions have a positive impact (Borghardt, 2022). Thus, future research on FALC should focus on how specific types of dependency relations affect comprehensibility. A more fine-grained analysis of the dependency relations would be interesting as the current analysis did not account for the numbers of dependencies per sentence.

In conclusion, ChatGPT and Le Chat produced target texts that are a good starting point, but post-editing is needed. Currently, these LLMs cannot replace the work of a human translator, although the human translator did not outperform the LLMs in each category. However, if they are seen as a tool to support the translation process, especially to save time, they have a lot of potential.

The validity of the results of this pilot study is limited by the rather small corpus and the fact that the qualitative analysis could not be carried out under the four-eye-principle. Therefore, future research will focus on enlarging the corpus and including other text types and domains.

Furthermore, it would be interesting to compare different prompts and strategies. Although recommendations for prompting like assigning a role were taken into account here, more iterations and few-shot in-context examples, as suggested by (Hui Jiao et al., 2024), were not tested.

References

Abdul Rauf, S., Ligozat, A.-L., Yvon, F., Illouz, G., & Hamon, T. (2020). Simplification automatique de texte dans un contexte de faibles ressources (Automatic Text Simplification : Approaching the Problem in Low Resource Settings for French). In C. Benzitoun, C. Braud, L. Huber, D. Langlois, S. Pogodalla, & S. Schneider (Eds.), *Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 2 : Traitement Automatique des Langues Naturelles* (pp. 332–341). ATALA. https://aclanthology.org/2020.jeptalnrecital-taln.33/

Anschütz, M., Oehms, J., Wimmer, T., Jezierski, B., & Groh, G. (2023). Language Models for German Text Simplification: Overcoming Parallel Data Scarcity through Style-specific Pre-training. *Findings of the Association for Computational Linguistics: ACL 2023*, 1147–1158. https://doi.org/10.18653/v1/2023.findings-acl.74

Baker, M. (1995). Corpora in Translation Studies. An Overview and Some Suggestions for Future Research. *Target. International Journal of Translation Studies*, *7*(2), 223–243. https://benjamins.com/online/target/articles/target.7.2.03bak

Bestgen, Y. (2024). Measuring Lexical Diversity in Texts: The Twofold Length Problem. *Language Learning*, *74*(3), 638–671. https://doi.org/10.1111/lang.12630

Borghardt, L. (2022). *Die Syntax von Leichter Sprache: Reduziert die Umformulierung in Einzelsätze die Komplexität? Eine fMRT-Studie zur Verarbeitung von Kausalsätzen* [, Johannes Gutenberg-Universität Mainz]. DataCite.

Braune, E. (2024, March 14). Le succès de Mistral AI exaspère l'UE! *The Conversation*. https://theconversation.com/le-succes-de-mistral-ai-exaspere-lue-225385

Brouwers, L., Bernhard, D., Ligozat, Anne-Laure, & François, T. (2014). Syntactic sentence simplification for French. In *EACL,* Gothenburg, Sweden.

Campesato, O. (2024). *Large Language Models. An Introduction. MLI Generative AI Series*. Mercury Learning and Information. https://doi.org/10.1515/9781501520587

Canut, E., Delahaie, J., & Husianycia, M. (2020). Vous avez dit FALC? Pour une adaptation linguistique des textes destinés aux migrants nouvellement arrivés. *Langage et societe*, *171*(3), 171–201. https://www.cairn.info/revue-langage-et-societe-2020-3-page-171.htm

Cardon, R., & Grabar, N. (2019). Parallel Sentence Retrieval From Comparable Corpora for Biomedical Text Simplification. In *Recent Advances in Natural Language Processing,* Varna, Bulgaria.

Cardon, R., & Grabar, N. (2020). French Biomedical Text Simplification: When Small and Precise Helps. In D. Scott, N. Bel, & C. Zong (Eds.), *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 710–716). International Committee on Computational Linguistics. https://doi.org/10.18653/v1/2020.coling-main.62

Charte d'accessibilité de la communication de l'État, 2022. https://www.gouvernement.fr/sites/default/files/contenu/piece-jointe/2023/03/sig_charte_accessibilite_v15_balise.pdf

Chehab, N., Holken, H., & Malgrange, M. (2019). *Rapport Final: Étude Recueil des besoins FALC*. http://51.91.138.70/simples/docs/SIMPLES_Etude_Recueil_desBesoins_FALC_HC.pdf

Christmann, U., & Groeben, N. (2019). Verständlichkeit: die psychologische Perspektive. In I. Rink & C. Maaß (Eds.), *Handbuch Barrierefreie Kommunikation* (pp. 123–145). Frank & Timme.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *Journal of Quantitative Linguistics*, *17*(2), 94–100. https://doi.org/10.1080/09296171003643098

Deilen, S., Hernández Garrido, S., Lapshinova-Koltunski, E., & Maaß, C. (2023). Using ChatGPT as a CAT tool in Easy Language translation. *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, 1–10. https://doi.org/10.26615/978-954-452-086-1_001

ELIS. (2024). *European Language Industry Survey 2024*. https://elis-survey.org/wp-content/uploads/2024/03/ELIS-2024-Report.pdf

Feng, Y., Qiang, J., Li, Y., Yuan, Y., & Zhu, Y. (2023). *Sentence Simplification via Large Language Models.* https://doi.org/10.48550/arXiv.2302.11957

François, T., Brouwers, L., Naets, H., & Fairon, C. (2014). AMesure: a readability formula for administrative texts (AMESURE: une plateforme de lisibilité pour les textes administratifs)[in French]. In *TALN,* Marseille. https://aclanthology.org/f14-2014.pdf

François, T., Müller, A., Rolin, E., & Norré, M. (2020). AMesure: A Web Platform to Assist the Clear Writing of Administrative Texts. In D. Wong & D. Kiela (Eds.), *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing: System Demonstrations* (pp. 1–7). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.aacl-demo.1

Gala, N., Tack, A., Javourey-Devet, L., François, T., & Ziegler, J. C. (2020). Alector: A Parallel Corpus of Simplified French Texts with Alignments of Misreadings by Poor and Dyslexic Readers. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 1353–1361). European Language Resources Association (ELRA). https://aclanthology.org/2020.lrec-1.169.pdf

Gao, Y., Wang, R., & Hou, F. (2024). How to Design Translation Prompts for ChatGPT: An Empirical Study. In R. Wang, Z. Wang, J. Liu, A. Del Bimbo, J. Zhou, A. Basu, & M. Xu (Eds.), *Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops* (pp. 1–7). ACM. https://doi.org/10.1145/3700410.3702123

Grabar, N., & Cardon, R. (2018). CLEAR – Simple Corpus for Medical French. In A. Jönsson, E. Rennes, H. Saggion, S. Stajner, & V. Yaneva (Eds.), *Proceedings of the 1st Workshop on Automatic Text Adaptation (ATA)* (pp. 3–9). Association for Computational Linguistics. https://doi.org/10.18653/v1/W18-7002

Hansen-Schirra, S., Bisang, W., Nagels, A., Gutermuth, S., Fuchs, J., Borghardt, L., Deilen, S., Gros, A.-K., Schiffl, L., & Sommer, J. (2020). Intralingual Translation into Easy Language - Or how to Reduce Cognitive Processing Costs. In S. Hansen-Schirra & C. Maaß (Eds.), *Easy--Plain--Accessible: Vol. 2. Easy Language Research: Text and User Perspectives* (pp. 197–225). Frank & Timme.

Hansen-Schirra, S., & Maaß, C. (2020). Easy Language, Plain Language, Easy Language Plus: Perspectives on Comprehensibility and Stigmatisation. In S. Hansen-Schirra & C. Maaß (Eds.), *Easy--Plain--Accessible: Vol. 2. Easy Language Research: Text and User Perspectives* (pp. 17–38). Frank & Timme.

Hansen-Schirra, S., Nitzke, J., Gutermuth, S., Maaß, C., & Rink, I. (2020). Technologies for the Translation of Specialised Texts into Easy Language. In S. Hansen-Schirra & C. Maaß (Eds.), *Easy--Plain--Accessible: Vol. 2. Easy Language Research: Text and User Perspectives* (pp. 99–127). Frank & Timme.

He, S. (2024, February 29). *Prompting ChatGPT for Translation: A Comparative Analysis of Translation Brief and Persona Prompts*. http://arxiv.org/pdf/2403.00127

Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., Kim, Y. J., Afify, M., & Awadalla, H. H. (2023, February 18). *How Good Are GPT Models at Machine Translation? A Comprehensive Evaluation*. http://arxiv.org/pdf/2302.09210v1

Hui Jiao, Bei Peng, Lu Zong, Xiaojun Zhang, & Xinwei Li (2024). Gradable ChatGPT Translation Evaluation. *Procesamiento del Lenguaje Natural*, *72*(0), 73–85. http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6579

Inclusion Europe. (2009). *L'information pour tous.: Règles européennes pour une information facile à lire et à comprendre*. UNAPEI.

Jiao, W., Wang, W., Huang, J., Wang, X., Shi, S., & Tu, Z. (2023). *Is ChatGPT A Good Translator? Yes With GPT-4 As The Engine.* https://doi.org/10.48550/arXiv.2301.08745

Kettunen, K. (2014). Can Type-Token Ratio be Used to Show Morphological Complexity of Languages? *Journal of Quantitative Linguistics*, *21*(3), 223–245. https://doi.org/10.1080/09296174.2014.911506

Kew, T., Chi, A., Vásquez-Rodríguez, L., Agrawal, S., Aumiller, D., Alva-Manchego, F., & Shardlow, M. (2023). *BLESS: Benchmarking Large Language Models on Sentence Simplification.* https://doi.org/10.48550/arXiv.2310.15773

Klöser, L., Beele, M., Schagen, J.-N., & Kraft, B. (2024). *German Text Simplification: Finetuning Large Language Models with Semi-Synthetic Data.* https://doi.org/10.48550/arXiv.2402.10675

Lindholm, C., & Vanhatalo, U. (Eds.). (2021a). *Easy - plain - accessible: vol. 8. Handbook of easy languages in Europe*. Frank & Timme. https://library.oapen.org/bitstream/id/1e86bbfc-0824-4040-80f9-d24f8a6d72d0/Handbook_of_Easy_Languages_in_Europe.pdf

Lindholm, C., & Vanhatalo, U. (2021b). Introduction. In C. Lindholm & U. Vanhatalo (Eds.), *Easy - plain - accessible: vol. 8. Handbook of easy languages in Europe* (pp. 11–26). Frank & Timme.

Loi n° 2005-102 du 11 février 2005 pour l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées, 2005. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000809647

Maaß, C. (2020). *Easy Language - Plain Language - Easy Language Plus*. Frank & Timme.

Maaß, C., & Rink, I. (2020). Scenarios for Easy Language Translation: How to Produce Accessible Content for Users with Diverse Needs. In S. Hansen-Schirra & C. Maaß (Eds.), *Easy--Plain--Accessible: Vol. 2. Easy Language Research: Text and User Perspectives* (pp. 41–56). Frank & Timme.

Marneffe, M.-C. de, Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal Dependencies. *Computational Linguistics*, 1–54. https://doi.org/10.1162/coli_a_00402

Ministère du Travail, de la Santé, des solidarités et des familles. (2025). *Accessibilité : une campagne de sensibilisation à destination des agents publics*. https://www.handicap.gouv.fr/accessibilite-une-campagne-de-sensibilisation-destination-des-agents-publics

Mistral AI. (n.d.). *Accompagner nos clients à la pointe.* Retrieved April 8, 2025, from https://mistral.ai/fr/customers

Ormaechea, L., & Tsourakis, N. (2023). Extracting Sentence Simplification Pairs from French Comparable Corpora Using a Two-Step Filtering Method. In H. Ghorbel, M. Sokhn, M. Cieliebak, M. Hürlimann, de Salis, Emmanuel, & J. Guerne (Chairs), *SwissText,* Neuchâtel, Switzerland. https://aclanthology.org/2023.swisstext-1.4/

Ormaechea, L., & Tsourakis, N. (2024). Automatic text simplification for French: model fine-tuning for simplicity assessment and simpler text generation. *International Journal of Speech Technology*, *27*(4), 957–976. https://doi.org/10.1007/s10772-024-10146-0

Parpan-Blaser, A., Girard-Groeber, S., Antener, G., Arn, C., Baumann, R., Caplazi, A., Carrer, L., Diacquenod, C., Lichtenauer, A., & Sterchi, A. (2021). Easy Language in Switzerland. In C. Lindholm & U. Vanhatalo (Eds.), *Easy - plain - accessible: vol. 8. Handbook of easy languages in Europe.* Frank & Timme.

Qi, P., Zhang, Y [Yuhao], Zhang, Y [Yuhui], Bolton, J., & Manning, C. D. (2020). Stanza: A Python Natural Language Processing Toolkit for Many Human Languages.

Qiang, J., Huang, M., Zhu, Y., Yuan, Y., Zhang, C., & Yu, K. (2025). *Redefining Simplicity: Benchmarking Large Language Models from Lexical to Document Simplification.* https://doi.org/10.48550/arXiv.2502.08281

Rink, I. (2019). Kommunikationsbarrieren. In I. Rink & C. Maaß (Eds.), *Handbuch Barrierefreie Kommunikation* (pp. 29–65). Frank & Timme. https://doi.org/10.25528/020

Rivas Ginel, M. I., & Moorkens, J. (2024). A year of ChatGPT: translators' attitudes and degree of adoption. *Tradumàtica Tecnologies De La Traducció*(22), 258–275. https://doi.org/10.5565/rev/tradumatica.369

Rodríguez Vázquez, S., Kaplan, A., Bouillon, P., Griebel, C., & Azari, R. (2022). La traduction automatique des textes faciles à lire et à comprendre (FALC) : une étude comparative. *Meta*, *67*(1), 18–49. https://doi.org/10.7202/1092189ar

Rolin, E., Langlois, Q., Watrin, P., & François, T. (2021). FrenLyS: A Tool for the Automatic Simplification of French General Language Texts. In *Proceedings of the Conference Recent Advances in Natural Language Processing - Deep Learning for Natural Language Processing Methods and Applications* (pp. 1196–1205). INCOMA Ltd. Shoumen, BULGARIA. https://doi.org/10.26615/978-954-452-072-4_135

Saggion, H. (2024). Artificial intelligence and natural language processing for easy-to-read texts. *Revista De Llengua I Dret*(82), 84–103. https://doi.org/10.58992/rld.i82.2024.4362

Seretan, V. (2012). Acquisition of Syntactic Simplification Rules for French. In N. Calzolari, K. Choukri, T. Declerck, M. Uğur Doğan, B. Maegaard, J. Mariani, . . . S. Piperidis (Chairs), *LREC,* Istanbul, Turkey. http://www.lrec-conf.org/proceedings/lrec2012/pdf/304_Paper.pdf

Siddharthan, A. (2014). A survey of research on text simplification. *ITL - International Journal of Applied Linguistics*, *165*(2), 259–298. https://doi.org/10.1075/itl.165.2.06sid

Todirascu, A., Wilkens, R., Rolin, E., François, T., Bernhard, D., & Gala, N. (2022). HECTOR: A Hybrid TExt SimplifiCation TOol for Raw Texts in French. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Thirteenth Language Resources and Evaulation Conference. LREC* (pp. 4620–4630). European Language Resources Association (ELRA). https://aclanthology.org/2022.lrec-1.493/

Vandeghinste, V., Müller, A., François, T., & Clercq, O. de. (2021). Easy Language in Belgium. In C. Lindholm & U. Vanhatalo (Eds.), *Easy - plain - accessible: vol. 8. Handbook of easy languages in Europe* (pp. 53–90). Frank & Timme.

Vilar, D., Freitag, M., Cherry, C., Luo, J., Ratnakar, V., & Foster, G. (2022). *Prompting PaLM for Translation: Assessing Strategies and Performance.* https://doi.org/10.48550/arXiv.2211.09102

Yamada, M. (2023). Optimizing Machine Translation through Prompt Engineering: An Investigation into ChatGPT's Customizability. In *Proceedings of Machine Translation Summit XIX, Vol. 2: Users Track,* Macau, China.

Zhang, X., Rajabi, N., Duh, K., & Koehn, P. (2023). Machine Translation with Large Language Models: Prompting, Few-shot Learning, and Fine-tuning with QLoRA. In P. Koehn, B. Haddow, T. Kocmi, & C. Monz (Eds.), *Proceedings of the Eighth Conference on Machine Translation* (pp. 468–481). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.wmt-1.43

Zhu, W., Liu, H., Dong, Q., Xu, J., Huang, S., Kong, L., Chen, J., & Li, L. (2024). Multilingual Machine Translation with Large Language Models: Empirical Results and Analysis. In K. Duh, H. Gomez, & S. Bethard (Eds.), *Findings of the Association for Computational Linguistics: NAACL 2024* (pp. 2765–2781). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-naacl.176

A Appendix: Prompts

Prompt 1 :
Tu es traductrice professionnelle. Tu fais des traductions intralinguales du français standard vers le FALC (français facile à lire et à comprendre). Le public cible a des difficultés de lecture. Le domaine de spécialité des textes originaux est la santé, plus précisément la santé sexuelle. Voici les principes de base du FALC :
Règles de rédaction :
Utiliser des phrases courtes (une seule idée par phrase).
Employer des mots simples et connus (éviter le jargon, les sigles et les abréviations).
Préférer la voix active (ex. : Marie ouvre la porte plutôt que La porte est ouverte par Marie).
Expliquer les mots compliqués si leur utilisation est indispensable.
Éviter les négations doubles (ex. : écrire C’est possible au lieu de Ce n’est pas impossible).
Faire des listes avec des puces pour organiser l'information.
Utiliser des exemples concrets pour illustrer une idée.
Mise en page et présentation :
Écrire en gros caractères (taille 14 minimum, en Arial ou Verdana).
Aérer le texte (un seul concept par ligne).
Utiliser des images ou pictogrammes pour illustrer les concepts importants.
Aligner le texte à gauche (éviter le texte justifié).
Mettre en gras les mots importants (éviter l’italique et le souligné).

Le FALC est souvent utilisé dans les documents administratifs, les brochures d'information et les sites web pour améliorer l'accessibilité. Traduit le texte suivant en FALC en appliquant les règles qui sont citées en haut et en rajoutant des explications des mots si tu le juges nécessaire. Il est important de conserver les informations clés du texte. Le texte cible doit être un texte en FALC, qui correspond aux règles. Voici le texte à traduire : […]

Prompt 2 :
Simplifie encore plus le texte, les informations clés doivent être conservées, mais le lexique et la syntaxe peuvent être simplifiés.

Prompt 3 :
Simplifie encore le texte en prenant en compte les règles du FALC, les informations clés doivent être conservées. Simplifie le lexique et la syntaxe et rajoute des explications si c'est nécessaire pour la compréhension.

B Appendix: List of Source Texts (ST)

	Subject	Author/Editor	Link
ST_1	Breast Cancer Screening	Institut National du cancer	https://www.crcdc-hdf.fr/wp-content/uploads/2023/03/Depliant-DOCS-2022_148x210-DEPSEIN21-BD-4.pdf
ST_2	Abortion	Ministère de la Santé et de la Prévention	https://ivg.gouv.fr/sites/ivg/files/2022-11/IVG%20Guide%20complet.pdf
ST_3	Mental Health	Ministère de la Santé et de la Solidarité	https://sante.gouv.fr/IMG/pdf/sante-mentale-guide-adultes.pdf
ST_4	Sexual Health, consent	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/TngXksBko3DWzyb
ST_5	Gender identity and sexual orientation	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/HpN9kCbb3C6pmHx
ST_6	Violence and sexual assault	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/8jiHJDXk9QeXFgp
ST_7	Contraceptives	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/TsGYWdBYEWsEnF2
ST_8	Abortion	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/nFCzomgFxfYErEP
ST_9	Morning-after pill	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/s54jBx3PQs3kREt
ST_10	Menstruation	Planning familial, Région Nouvelle Aquitaine	https://cloud6.zourit.net/index.php/s/TtCzKTtcRwWjy9k
ST_11	Sexually transmitted diseases	Planning familial, Région Nouvelle Aquitaine	https://www.calameo.com/read/0075046587a946b2beb4c
ST_12	Sexually transmitted diseases	Planning familial des Pyrénées Atlantiques	https://www.tonplanatoi.fr/uploads/images/FALC_Plaquette_Planning_Familial_PAU_2024-1(1).pdf
ST_13	Contraceptives	Planning familial des Pyrénées Atlantiques	https://www.tonplanatoi.fr/uploads/images/FALC_Plaquette_Planning_Familial_PAU_2024-1(1).pdf
ST_14	Abortion	Planning familial des Pyrénées Atlantiques	https://www.tonplanatoi.fr/uploads/images/FALC_Plaquette_Planning_Familial_PAU_2024-1(1).pdf
ST_15	Violence	Planning familial des Pyrénées Atlantiques	https://www.tonplanatoi.fr/uploads/images/FALC_Plaquette_Planning_Familial_PAU_2024-1(1).pdf