

Towards Clear, Simple, and Inclusive Legal Language

— Exploring Generative Artificial Intelligence (GAI) models as content rephrasing assistants in multilingual settings

*Maria Zimina-Poirot, Christopher Gledhill and Manon Bouyé**

Abstract

This study explores the potential for GAI language models and Natural Language Processing (NLP) tools to be used as writing aids in the transformation of expert content into clear, simple and inclusive language. In particular, we focus on the feasibility of fully automating the reformulation of legal texts, and the challenges of setting standards for automatically generating content in multilingual settings. While automatic assistants can already produce effective summaries and simplified versions of complex legal texts, we suggest that they lack the functional awareness and metalinguistic reflexivity that are particular to human review and revision. We illustrate this by examining a series of authentic examples. First, we look at tools like ChatGPT which can be configured to summarise complex legal documents and generate texts aligned using the guidelines for Plain Language (PL) and Easy-to-Read (E2R). We give examples of linguistic features, outputs from NLP models, and GAI-generated content evaluation. Our initial observations suggest that while automatic assistants can produce greatly simplified texts, they are unable to take into account legal implications and contextual reasoning. Then we look at the issue of prescriptive guidelines in the light of the new paradigm of content management. Comparing examples of official discourse and their equivalents ‘translated’ into E2R, we find that the reformulations involved reflect the very different functional requirements of such radically simplified texts. This observation leads us to stress the importance of human oversight in auto-generated content, as well as the need for guidelines, integrated revision tools, transparency, user training, and continuous monitoring.

Submitted: 23.02.2025, accepted: 23.02.2025, published online: 20.05.2025

*Maria Zimina-Poirot: Université Paris Cité, maria.zimina-poirot@u-paris.fr; Christopher Gledhill: Université Paris Cité, christopher.gledhill@u-paris.fr; Manon Bouyé: Université Jean Moulin, manon.bouye@univ-lyon3.fr.

Keywords

computer-mediated communication (CMC), easy-to-read (E2R), generative artificial intelligence (GAI), legal discourse, natural language processing (NLP), plain language (PL), recontextualisation, training for language professionals.

1. Introduction

The aim of this paper is to explore how Generative Artificial Intelligence (GAI) and Natural Language Processing (NLP) are being used as editorial tools in the production of Plain Language (PL) and other adapted versions of administrative and legal texts. The new technologies raise a number of questions for language analysts, as well as for content providers and end-users. For example, given that non-expert users can now readily use the currently available technology to produce working summaries of expert texts, to what extent is the output of automatic text reformulation ‘formally adequate’ (providing coherent, fluent reformulations based on the original text), and to what extent is the output of these tools ‘functionally adequate’ (providing relevant, user-oriented adaptations)?

To answer these questions systematically would require significant data analysis, preferably with experimental protocols. There now exists a body of research on the quantitative analysis of legal simplification of legal discourse (for example: Kline, 2023; Mamac, 2024), and we have ourselves conducted such analysis (Bouyé, 2022; Bouyé & Gledhill, 2024). However, the aim of this paper is not to provide quantitative analysis, but rather to raise a number of questions and to explore some basic principles which we believe will be worth examining in subsequent work. It should be added that we focus here primarily on texts produced and reformulated in English, but we are also interested in simplification in multilingual settings (explanatory texts produced by multilingual organisations such as the EU, or adaptations of texts into Plain Language produced by users of English as an additional language). As we explain below, these questions are relevant because we are teachers/researchers engaged in graduate programmes that involve the training of specialised translators. One of the key competencies for these students is that they should be familiar with the concept and practice of writing in clear, simple and inclusive language (PL, as well as Easy-to-Read, controlled languages etc.) and should be familiar with the technologies that have been developed recently to enable translators and technical communicators to produce simplified versions of the texts for their clients.

The first issue we raise is whether automatic text reformulation is **formally and functionally adequate**. We address this question in the following discussion by looking at the simplification of administrative and legal texts conducted by tools such as ChatGPT. As we mention below, ‘simplification’ can refer to the formal transformation of an expert-to-expert text into a more restricted variety (Plain Language, Easy-to-Read, or some other). But the notion is clearly more subtle than mere textual manipulation; it also requires some representation of how interpersonal relations are to be realised in the interaction (for example, what are the roles of participants involved, and what rhetorical strategies do they wish to adopt? Expert vs. non-expert? Mediator vs. petitioner? etc.). As we argue below, it appears that automatic systems are good at implementing textual

reformulations, but in comparison with explanatory texts produced by humans, the functional dimension appears to be largely lacking.

The second principle we wish to address here concerns the concept of **recontextualisation**, especially as it has been adopted by analysts working within the framework of Systemic Functional linguistics, SFL (Van Leeuwen, 1993; Linell, 1998). Recontextualisation refers to the process by which social practices are assigned new meaning when transferred across cultural and linguistic contexts. Key to this is the realisation that the function of a text or dialogue is not to transmit information, but rather to perform social interaction, with meaning being seen as continuously ‘re-negotiated’ as discourse unfolds (Rock, 2007). Three sub-types of recontextualisation have been proposed (Linell, 1998): i) ‘intratextual’ recontextualisation, that is to say the recreation of meaning within the same on-going text or discourse, ii) ‘intertextual’ recontextualisation, involving rewording and cross-references between texts (of the same language), and iii) ‘interdiscursive’ reformulation which, as we see below, concerns the recreation of meaning across and between different genres and text types.

How does recontextualisation relate to simplification and the reformulation of texts in Plain Language? The SFL approach is not to see summarisation and simplification as operations that are applied to an original text, but rather as differing functional goals which are realised by a new context of communication, whose frame of reference is defined by the original source text (or speech). In the case of summarisation, the new context allows for the target text to omit or modify certain details from the source material, at the same time as maintaining the original text function (in the case of legal definitions, for example, this function should be preserved). Conversely, simplification involves a reconfiguration of the functional parameters of the text (in many cases, the end-user is asking the text to ‘do something different’: for example, they are looking for advice, asking questions, etc.). In the context of the law, the recontextualisation of legal knowledge is essential for laypeople to have access to justice and to know their rights. Two main functions can be identified for recontextualised legal texts. The first is to transform mainly descriptive or argumentative legal texts into informational texts, which transfer information about the legal domain, its actors, terminology and latest developments, usually in a specific national or multinational context (Bouyé, 2022: 71). The other main objective of the recontextualisation of specialised legal knowledge is to produce instructional or procedural texts (Adam, 2001), which not only popularise theoretical legal knowledge, but also transfer practical and procedural “know-how” so that laypeople can take informed decisions in specific legal or administrative situations (Prite, 2018). As we suggest below, GAI and other tools are good at producing text summaries, but they do yet not address interdiscursive reformulation; that is to say they cannot fulfill the functional requirements of simplification that can only be determined when one has access to the end-users’ goals and expectations. It is also important to include the concept of multimodality in any discussion of recontextualisation: as we see below, the more radical forms of text reformulation, as implemented by Easy-to-Read, involve

non-linguistic features (images, text formatting) that nevertheless have a significant impact on the lexicogrammatical characteristics of simplified texts.

The third issue we wish to address here is the paradigm shift which we perceive in the relationship between the producers and consumers of texts, especially in terms of how the responsibility for ‘policing’ language is shifting from human institutions and experts to **third party language mediators**. There has long been a demand for popularised or simplified versions of administrative and legal content, including the communications produced by public-facing government departments, advice provided by charity organisations and the like (Krieg-Planque, 2020).

Traditionally, the production of this material was in the hands of the original authors, communication departments or expert writers (and in the case of legal texts: ‘jurilinguists’ (Baumert, Geslin, Roussel & Schott, 2021)). We suggest here that the advent of GAI and NLP tools poses a challenge for the traditional model of text reformulation, not only in terms of how content is generated and mediated, but also how this content is ‘policed’ or mediated by third parties. This is based on the assumption that automatic reformulation will be widespread in the future, and that much of the linguistic content processed by these tools will not be created or controlled directly by the originators of the information, but will be generated ‘on the fly’, for example by chatbots and other dialogue simulators. For example, will it be possible to distinguish between content that has been validated by human revisers, as opposed to content processed entirely by the AI? And to what extent is it possible for the original authors (whether content providers or experts, etc.) to control the output, that is to say to impose their own linguistic recommendations or editorial preferences on the reformulated content? Similarly, to what extent is it possible for either the content providers or the end-users to ensure that the output they are consuming does not contain bias (discrimination relating to race, gender, mental and physical abilities, etc., as discussed in Kline (2023) and Helm, Bella, Koch & Giunchiglia (2024)?

The following paper is divided into five main sections, plus a conclusion. In Section two, we examine some of the currently available NLP tools that have been designed to generate legal summaries, with a particular emphasis on the need for training in how to operate these systems. In Section three, we examine the impact of GAI and NLP tools on the future of two main prescriptive instruments used by language professionals to ‘police’ language (style guides and language policies). We then look at a sample of extracts produced using the Easy-to-Read guidelines, with the aim of showing how the lexicogrammatical features of these texts reflect the functional requirements of this very specific form of simplification. Section four then addresses the question of plain language as a competency: here we emphasise the highly interdependent relationship between technological competencies on the one hand, and linguistic experience and sensitivity. In Section 5, we raise the question of intersubjective meta-reflexivity, asking ChapGPT to examine how it sees its own relationship with language professionals. In the concluding Section, we argue that the role of human experts is essential in the

development of AI-assisted translation and reformulation tools, in order to ensure the quality, clarity, and reliability of automatically generated legal texts.

2. Generative AI in Multilingual Content Creation and Translation Education

Recent advances in GAI technologies, particularly those involving Large Language Models (LLMs), are revolutionising translation and multilingual communication. These models surpass traditional Neural Machine Translation (NMT) engines in certain tasks, as demonstrated by a 2023 Lionbridge evaluation (Lionbridge, 2023). LLMs are developed using sophisticated machine learning techniques and comprise a vast number of parameters (Roffo, 2024). OpenAI's large multimodal models (accepting text or image input and outputting text), such as GPT-4, represent significant progress in this field. This type of model has been trained on an extensive dataset comprising a diverse range of textual sources, and is designed to enhance performance, accuracy, and contextual understanding. One of its principal characteristics is its capacity to process and generate text in a multitude of languages, incorporating a wide range of references up to the knowledge cut-off date, depending on the temporal framework of the training data (Zhang, Dong, Xiao & Oyamada, 2024). This multilingual capability is supported by linguistic transfer, which enables the model to draw on semantic knowledge from a range of languages, thereby enhancing its comprehension and output.

In translation education, integrating LLMs into teaching workflows presents both opportunities and challenges. It requires balancing technological advances with the imperative to maintain learner autonomy and develop core translation competencies (Zimina-Poirot, forthcoming). We conclude that Generative AI tools can (or at least should) be positioned as valuable aids rather than substitutes for human translators. Learners must retain control over the translation process, critically evaluating the output of automated systems and developing their ability to assess the quality and appropriateness of translations (Gledhill & Zimina-Poirot, 2022; Zimina-Poirot & Gledhill, 2023). In addition, the effectiveness of human-machine collaboration hinges on learners acquiring the skills to provide clear and precise instructions to LLMs (Kantharaja, Srinivasalu, Rajashekara, Vishvanath & Sanjeev, 2024).

These principles form the foundation of our approach to teaching plain and inclusive language in a translation project conducted as part of a European Master's in Translation (EMT) programme: ILTS, Université Paris Cité.¹ One of the key modules of this

¹ *Master Traduction, interprétation, Industrie de la langue et traduction spécialisée (ILTS)*. Available at odf.u-paris.fr/fr/offre-de-formation/master-XB/arts-lettres-langues-ALL/traduction-interpretation-K6JMSAFS/master-traduction-interpretation-parcours-industrie-de-la-langue-et-traduction-specialisee-JRQNFV5Z.html (accessed 07 May 2025).

project involves students translating a French website into English. Between 2020 and 2024, we collaborated with the French Federation of People with Diabetes (FFD),² which, like many similar charities, develops policy initiatives aimed at challenging misconceptions about diabetes. Engaging with such ‘disruptive discourse’ introduces unique challenges related to teaching objectives, technology integration, and the broader context of translation practices. Through this collaboration, we have identified several pedagogical and practical challenges, as well as solutions for effectively incorporating AI tools into the translation process (Gledhill & Zimina 2019; Gledhill & Zimina-Poirot, 2022). In this perspective, teaching plain and inclusive language requires a shift from detailed prescriptive methods to fostering students’ critical engagement. We encourage learners to develop concise editorial guidelines, enhancing their ability to make informed decisions (Zimina-Poirot & Gledhill, 2023). AI tools can help learners to analyse and refine drafts, as well as strengthen their editorial and critical thinking skills. Integrating generative AI tools such as ChatGPT into translation education enhances the learning experience by improving students’ editorial skills and their ability to critically manage machine-generated translations.

By addressing the challenges of learner autonomy, revision workflows and real-world client expectations, it is increasingly possible to equip students with the necessary tools and frameworks to navigate the complexities of multilingual translation (Zimina-Poirot, forthcoming). Through corpus analysis and reflection, these tools, when thoughtfully integrated, can support both the practical and theoretical aspects of translation training and prepare learners for professional demands.

An important area of research in this context is to raise students’ awareness of the strengths and weaknesses of generative AI tools and AI-assisted features, including the potential biases and inconsistencies they may encounter when using these advanced technologies such as ChatGPT-4.

The conversational abilities of ChatGPT-4 are a result of reinforcement learning from human feedback (Liu, 2024), where the model is trained using techniques such as annotation, rewards, and penalties. This training allows it to better align with user intent and maintain contextual relevance during interactions. Although ChatGPT-4 is not updated continuously, its conversational performance improves over time through user interactions and feedback, contributing to its adaptive learning.

When designing specialised applications, such as legal assistants, ChatGPT-4 can be customised using templates provided by tools like Keywords Everywhere.³ These templates specify key elements within prompts, such as clear instructions, relevant context, input data, and the desired output format. By clearly outlining these elements, the model can be guided to perform specific tasks, such as generating legal briefs,

² Available at federationdesdiabetiques.org/en/who-we-are (accessed 07 May 2025).

³ Available at keywordseverywhere.com/ (accessed 07 May 2025).

summaries, or reports, tailored to meet the unique requirements of various legal scenarios. Figure 1 illustrates the main elements included in a prompt:

- (1) Instructions: Clear and concise directions on what the legal assistant should do. This includes specific tasks, guidelines, or actions to be taken.
- (2) Context: Background information relevant to the task or scenario. This helps the legal assistant understand the environment or situation in which it will operate.
- (3) Input Data: The information or data that the legal assistant will work with. This could include legal documents, case details, client information, or any other relevant data.
- (4) Output Format: The format in which the results or responses should be presented. This could involve structured reports, summaries, legal briefs, or any other specified format.

Following this approach, prompts can be crafted to guide an AI-powered legal assistant in performing various professional roles (see Figure 1).

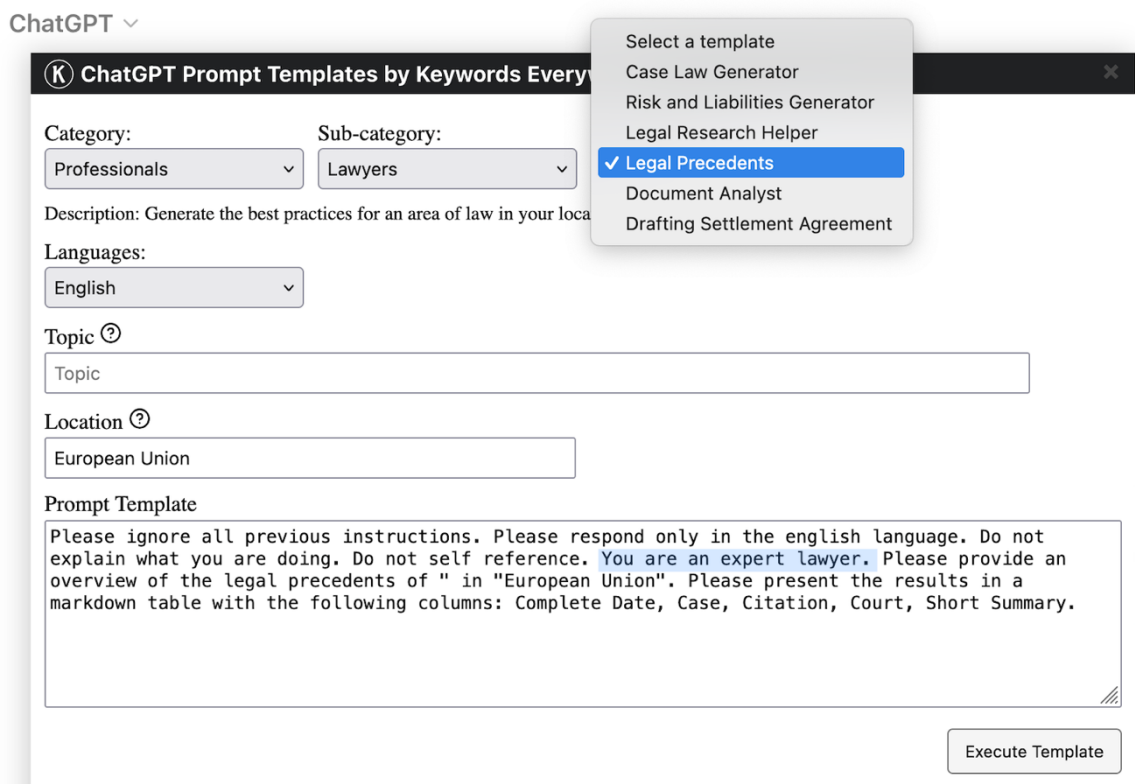


Figure 1: Example of ChatGPT Prompt Template by Keywords Everywhere containing the following elements: Instructions, Context, Input data, Output format (available in more than 30 languages).

Despite recent advances, LLM systems like ChatGPT still face several challenges, such as bias in training data, limitations in maintaining conversational memory, and high development costs (Kshetri, 2024; Martínez, Watson, Reviriego, Hernández, Juárez & Sarkar, 2023). Although automation can assist with certain aspects of legal text rephrasing, it is unlikely to replace nuanced understanding and expertise of human legal professionals, as demonstrated by the example in Table 1.

Alternative initiatives are being developed, which combine specialised artificial intelligence and natural language processing (NLP) models to develop automated assistants that facilitate enhanced access to legal texts. For instance, Legalese Decoder⁴ simplifies complex legal documents by transforming them into more comprehensible language, but may sacrifice the precision, formality, and technical detail required in professional or legal contexts (see the example on Figure 2). The EU-funded project Easy Reading is developing assistive software designed for users with cognitive impairments (Edler, 2020).⁵ Although the tool simplifies complex legal documents, it is evident that it still generates some errors and complexity that may impede full clarity for people with cognitive impairments (see the example in Figure 3).

⁴ Available at legalesedecoder.com (accessed 07 May 2025).

⁵ Available at easyreading.eu and easyreading.eu/downloads (accessed 07 May 2025).

Table 1: Plain Language: a comparison of human rephrasing and ChatGPT-powered AI rephrasing






Source text (Equality Act 2010)	Human rephrasing (Citizens Advice Bureau 2018)	GAI rephrasing (ChatGPT)
If the protected characteristic is age, A does not discriminate against B if A can show A's treatment of B to be a proportionate means of achieving a legitimate aim.	Sometimes it's not direct discrimination if someone treats you differently because of age. The law says it's possible to justify direct age discrimination, if there's a good enough reason for treating you differently. Example Your GP practice only provides free flu jabs to people over 65. This is because older people are more vulnerable to the flu and are more likely to be seriously ill, if they get the virus. This is a good enough reason for treating people differently based on age.	If someone is treating someone else differently because of their age, it is not considered discrimination if the person doing the treating can prove that they had a good reason for doing so and that their actions were fair and reasonable.

Note: Functionally speaking, the AI version is a re-definition of the original text, while the human rephrasing is action-oriented, involving instructional functions. In short, the human rephrasing uses more straightforward language, using concretisation strategies (Turnbull, 2018) in the form of an example situation to explain the reasoning behind the practice in concrete, familiar terms for lay readers.

These emerging trends, including the developments of AI-driven rephrasing assistants and NLP platforms, have resulted in a rapid increase in the number of available tools and features, as well as the volume of automatically generated text (Kilichev, Abbasova, Gafurova, Dadabayeva, Nurmatova & Fazildinova, 2024). These computer-assisted mediation tools are designed to enhance accessibility, comprehensibility, and relevance of textual content for users, while also improving productivity in information processing. However, they also raise important questions regarding the standards to be followed in this rapidly evolving field (O'Brien, 2024).

If experts are to employ artificial intelligence as a tool, they must also accept responsibility for monitoring and revising the outputs generated by GAI models in order to guarantee the quality of the content produced. This process can be time-consuming and requires significant human resources, which may explain the presence of errors in many multilingual documents generated through automated processes (see Section 3). In this context, questions arise regarding the standards to be followed (guidelines or norms that should govern how machines communicate with humans) and the paradigm to be adopted in human-machine interaction.

The challenge lies in finding a balance between leveraging GAI capabilities and maintaining human oversight to ensure both communicative efficiency and accuracy in legal contexts. In this regard, linguistic standards, specifications, and style guides provide an interesting basis for practical application (Gledhill & Zimina-Poirot, 2022). We examine issues relating to the use and acquisition of these guidelines in the following Section.

Free Legal Document To Plain English Translator - Instantly Interpret Free: Legalese Decoder - AI Lawyer Translate Legal Docs To Plain English

~2 minutes

Paste your legalese jargon in the first box, press Decode

Free access near capacity due to popularity 😊 [Subscribe](#) for FULL PAGE uploads.

Legal Document

The appropriations required for human resources and other expenditure of an administrative nature will be met by appropriations from the DG that are already assigned to management of the action and/or have been redeployed within the DG, together if necessary with any additional allocation which may be granted to the managing DG under the annual allocation procedure and in the light of budgetary constraints.

The money needed to pay for staff and other costs related to running the office will come from the budget of the department that is responsible for overseeing the project. This money may already be set aside for managing the project, or it may be moved from other parts of the department's budget. If needed, extra money may also be given to the department to help cover these costs, depending on the yearly budget process and how much money is available.

Figure 2: Legalese Decoder: rephrasing DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Brussels, 6.9.2023. COM(2023) 512 final. 2023/0311 (COD)

Note: The rephrased version appears to be clearer and more accessible. However, the reformulated text may lose some of the precision, formality, and detail required to understand formal administrative and financial documents. This highlights the trade-off between clearer communication and the loss of referential precision, which underscores the tendency of GAI reformulation to overgeneralise based on the input context or patterns in the pre-training corpus (Chang & Bergen, 2024).

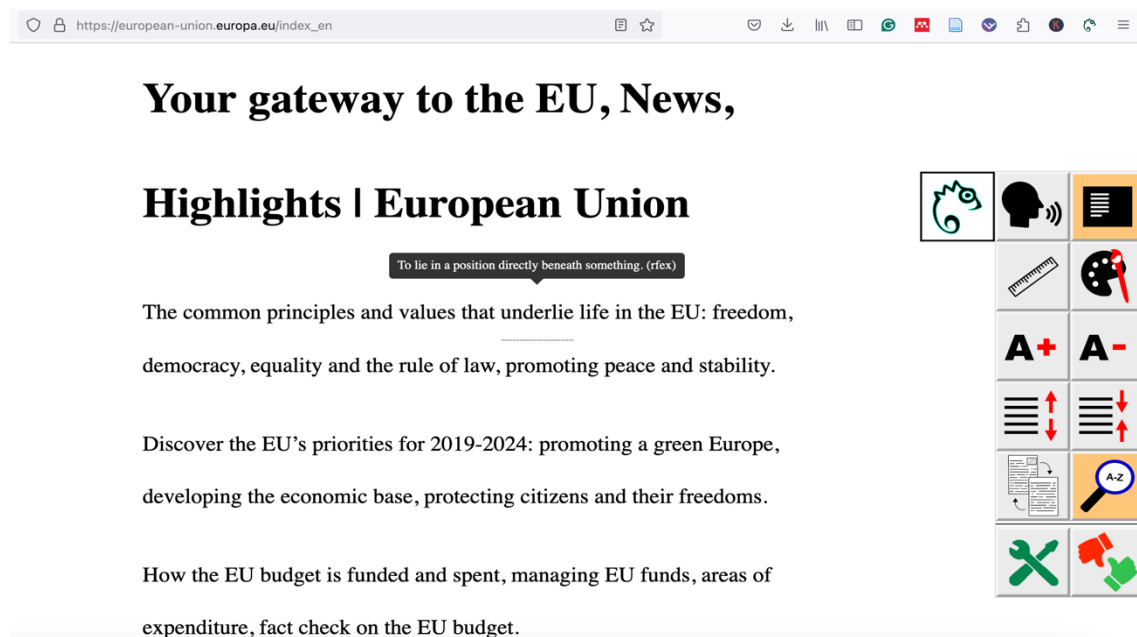


Figure 3: Easy Reading Program. Simplifying an official European Union website

Note: The combination of automatic rephrasing, text size adjustment, contrast control and text-to-speech functionality is designed to enhance the user experience. The rephrasing tool provides simplified definitions for complex words or phrases, such as the tooltip explaining the word “underlie”. However, this definition of “underlie” does not capture the intended meaning within the context, which could lead to misunderstanding.

3. Plain and Easy-to-Read Legal Writing: Definition and Institutional Framework

As explained in the introduction, the reformulation of specialised legal texts for laypeople is crucial for them to be able to know their rights, as citizens, and to make informed decisions, as customers, when they are in contact with the legal or judicial system (Turnbull, 2018; Preite, 2018). Before presenting some examples of legal reformulation using GAI tools, it is important to consider the linguistic guidelines on drafting standards that have been advocated and used by institutions to enhance the readability and accessibility of legal texts for different types of public.

Initiatives to simplify legal language have taken different forms and been adopted in various degrees over the years. The Plain Language Movement, which started in the 1970s and 1980s in the United Kingdom, United States and Australia, has had a lasting influence on drafting policies around the world. Plain Language (PL) can be defined as a set of linguistic and stylistic guidelines whose aim is to establish clear communication in administrative and legal contexts. The most frequent recommendations include syntactic reformulations, such as avoiding the passive, preferring short sentences, using verb forms rather than noun forms, as well as lexical replacements, notably avoiding

technical terms and formal lexis, often referred to as ‘difficult’ words. In terms of pragmatic recommendations, PL guidelines advocate user-focus and a clear layout. There has been considerable amount of research on PL, including attempts to measure the degree of adherence to PL standards in medical communication (Gledhill, Martikainen, Mestivier & Zimina-Poirot, 2019) and in the legal field (Bouyé, 2022).

In the European Union, the issue of the simplification of legal language was highlighted by the ‘Fight the FOG’ campaign, launched by translators in 2010 to encourage plain writing in various EU publications. Following this movement, the European Commission also published recommendations to expert drafters and translators to use PL, in the style guide *How To Write Clearly* (European Union, 2012), available in the 23 official languages of the European Union.

While Plain Language can be targeted at both expert-to-expert communication and non-expert settings, other forms of interdiscursive reformulation such as Easy-to-Read (E2R), known as *Facile à lire et à comprendre* in French (FALC)⁶ and *Leichte Sprache*⁷ in German, are aimed at specific segments of the population. Originally developed by and for people with cognitive impairments, E2R is now also used for second-language speakers or people with reading difficulties (Gangloff, 2015). Compared with PL, the E2R guidelines place strong emphasis on the use of multimodality, particularly images, to replace or illustrate written content. Recommendations also include breaking down information into one-line phrases, sentences with single clauses and “easy” vocabulary, prohibiting the use of foreign words or specialised terms, etc.

Several empirical studies have explored the actual uptake of E2R texts. For instance, research using eye-tracking technology (Hansen-Schirra, Bisang, Nagels, Gutermuth, Fuchs, Borghardt, Deilen, Gros, Schiffel & Sommer, 2020) has demonstrated that texts adapted according to E2R recommendations are read more quickly and are therefore more legible than non-adapted texts. This is a significant finding, since the textual reformulations required by the E2R guidelines are so far-reaching that it is clearly important to establish the empirical foundations of E2R before applying them more widely.

As mentioned above, E2R and FALC involve a number of similarities with the principles of clear style and Plain Language. However, these schemes also employ distinct strategies, particularly at the text level, which include radical changes in typography and format, as well as adjustments to text size and length (see Figure 3).

Cognitive linguists thus consider that there is a continuum of linguistic complexity, which ranges from highly complex specialised language to standard language (at one end of the scale), to PL and then finally to E2R (at the other end of the scale) (Hansen-Schirra & Maaß, 2020: 18).⁸

⁶ Available at falc.unapei.org/ (accessed 07 May 2025).

⁷ Available at leichte-sprache.de/leichte-sprache/was-ist-leichte-sprache (accessed 07 May 2025).

⁸ For a more detailed analysis of Easy Language and language policies in Europe, see Hansen-Schirra & Maaß (2020).

In the EU, E2R has been advocated by the organisation Inclusion Europe⁹ since its creation in 2009. E2R standards have for example been used to rewrite and simplify texts such as the Universal Declaration of Human Rights (Bock, 2018). Some pages of the European Commission's website also have an E2R version.¹⁰ This is in line with EU policies that are centered around the notions of inclusivity, including the 2019 Directive (UE) 2019/882 on the accessibility requirements for products and services, which requires products and services to be accessible for people with disabilities, and the publication of the European Disability Strategy in 2021. More recently, in 2023, the EU proposed a directive for a European Disability Card which has been widely publicised in the EU's general communications and in E2R as well as equivalent texts in the EU's other official languages.¹¹ In this context of accrued effort towards inclusivity, the issue of making legal and institutional information accessible appears crucial, which is why we have chosen to explore the potentials and challenges of applying E2R standards using GAI. Efforts towards more inclusive language have notably been put forth using style guides, which we turn to in the following Section.

3.1 Style Guides and the 'Gatekeepers of Language'

One of the key concepts traditionally presented to undergraduate students of English linguistics is the distinction between prescriptivism (seeking to formally fix language norms, by way of normative grammars, language academies, etc.) and descriptivism (recording and explaining language use, in the form of corpus-based dictionaries, etc.) (Aitchison, 1991; Pinker, 2015). We would argue that, with the advent of GAI and large language models, this distinction may be disappearing.

Ironically, it is the descriptive approach itself which led to most of the changes that we are witnessing (computational linguistics, etc.). It is now the case that descriptive linguists, who once relied on human-derived judgments and data, must now also analyse output that has been influenced by non-human language models. This leads to the need to distinguish between human-generated content and computer-mediated communication, with the implication that much of what we consider to be 'authentic' usage has been homogenised and filtered by technology. In this respect, using content labelling (or information about information) can help clarify the responsibilities and expectations for content creators and platforms in regulated environments (Morrow, Swire-Thompson, Polny, Kopec & Wihbey, 2022).

As far as the prescriptive approach is concerned, the use of GAI means that there is on the one hand huge scope for intervention, since the technology enables a degree of

⁹ Available at inclusion-europe.eu/ (accessed 07 May 2025).

¹⁰ For example: european-union.europa.eu/easy-read_en (accessed 07 May 2025).

¹¹ Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing the European Disability Card and the European Parking Card for persons with disabilities. COM/2023/512 final.

control over language that was simply not available in the past (moving beyond mere spelling and grammar checkers towards the potential imposition of comprehensive, automatically-generated style guides). On the other hand, it is easy to see how the role of humans as ‘gatekeepers’ or the ‘final arbiters’ of language norms may be challenged: as we see below, it is difficult for language reformers to impose their choices in the face of massive data sets that can now be used increasingly as the basis for any statistically derived norm.

However, let’s assume that the traditional distinction between the prescriptive and descriptive approaches is still relevant, and that human agents (at least for now) remain in control of language norms. From this perspective, both Plain Language (as a political movement) and Easy-to-Read (as a form of simplified language) firmly belong to the prescriptive approach and contribute to a renewed interest in language reform. In the English-speaking world, this is reflected not only in heightened public expectations for the use of Plain Language (evidenced by laws such as the *Plain Writing Act* of 2010), but also by the increasing use of inclusive style guides and other statements which attempt to modify the norms of English in order to reflect shifting social norms (explicitly signaling preferred gender and pronouns, using first-person language, etc.).

The prescriptive approach is often derided by academic linguists (e.g., cf. Pinker’s discussion of the ‘language mavens’, 1995), and for good reason: issues like arbitrary declarations and lack of evidence undermine its credibility. However, observers in the field of language planning (Lo Bianco, 2010) recognise that it is the responsibility of language gatekeepers (public authorities, publishers, educators, etc.) to make political decisions regarding which language forms should be codified (status planning, corpus planning, etc.).

In the context of plain and inclusive language, there have traditionally been two primary sources for language guidelines:

- 1) Style guides, published by official bodies at national or international levels, and
- 2) Position statements on language, issued by specialised organisations (charities, non-governmental organisations, etc.).

Table 2 presents examples of both types of prescriptive guidelines.

Table 2: Two prescriptive instruments on language

Source of authority	Examples of prescriptive guidelines
Style Guides	American Psychological Association (APA) 2021. <i>Inclusive Language Guide</i> .
	European Commission (DGT-EU) 2024. <i>English Style Guide. A handbook for authors and translators</i> .
	National Center on Disability and Journalism (NCDJ). <i>Disability Language Style Guide</i> .
Position Statements on Language	Diabetes Australia. 2021. <i>Our language matters: Improving communication with and about people with diabetes</i> .
	The Diversity Movement. 2020. <i>Say This, Not That. Activating Workplace Diversity Through Inclusive Language Practice</i> .
	LGBT+ Equity Center. 2019. <i>University of Maryland Policy on Inclusive Communication</i> .

The traditional style guide (or style manual) differs from other prescriptive documents, in that it attempts to encompass all of the editorial features that an organisation may require for its communicative needs. Unfortunately, there have only been a limited number of studies examining style guides from an empirical linguistic perspective (Crossley, Louwerse, McCarthy & McNamara, 2007). One example of this is Svoboda (2013), who examined the degree of compliance of EU websites with the European Union Interinstitutional style guide in a variety of languages.¹² He found that the guidelines are broadly respected when it comes to terminological and typographical practice, but there are also wide disparities in terms of content and style. Similarly, Bouyé (2022) has surveyed 15 style guides in English and French, finding that they offer consistent advice in terms of low-level features of language (terminology, passive voice, etc.). However, the recommendations become more inconsistent and ineffective at higher levels (sentence length, paragraph structure).

Style guides serve as important symbolic declarations of prescriptive language use. For example, the English Style Guide, published by the European Commission Directorate-General for Translation, stipulates that EU English should align with the varieties used in Ireland or Great Britain.¹³ While this is a useful policy statement, it is difficult to enforce practically. The guide addresses minor differences (like “-is” vs. “-iz-”) but lacks guidance on other, more profound distinctions between North American and European English (increasingly seen as parallel regional varieties). As a result, both professional and trainee translators often rely on the default settings of language tools, without much awareness of these nuances.

¹² Publications Office of the European Union, Interinstitutional style guide, Publications Office of the European Union, 2022. Available at data.europa.eu/doi/10.2830/215072 (accesses 07 May 2025).

¹³ Available at commission.europa.eu/resources-partners (accesses 07 May 2025).

Let us turn now to the second source of authority: position statements or policy statements on language. These documents are published by a variety of institutions, from governments to pressure groups, with the specific aims of (explicitly) influencing language practice in their respective fields and (implicitly) signaling the values of the institution. Although some documents may resemble declarations of intent¹⁴, others set out more detailed guidelines, together with justifications. Among the examples listed in Table 2, we would highlight the *Language Matters* campaign. The aim of *Language Matters* is to change the way people talk and write about diabetes, both among the general public and specialists. *Language Matters* originated as a set of guidelines advocated by a group of Australian specialists in the field (Speight, Skinner, Dunning, Black, Kilov, Lee, Scibilia & Johnson, 2021), and has subsequently spread internationally, as can be seen in documents such as *Le Pouvoir du Langage: langage et diabète* (published by a French association *Glucose Toujours*)¹⁵, and so on.

The conceptual scope of *Language Matters* is ambitious: it is not simply a question of changing a few technical terms, but of reorienting the entire phraseology of the field. To give just one example, the authors propose replacing the formulation “patient suffering from diabetes” by a seemingly more positive construction “person living with diabetes”. As noted by Froeliger, Gledhill, and Zimina-Poirot (2023), this type of reorientation not only carries symbolic significance by shifting the perspective from ‘third-party’ to ‘person-first’, but also has practical consequences. It requires extensive reformulations across the organisation’s textual output, including changes to the name of the organisation. For example, in a translation project conducted as part of the ILTS European Master’s in Translation (EMT),¹⁶ the M2 ILTS programme we mentioned above, M2 students had to translate the name *Fédération Française des Diabétiques* into English, which finally became “The French Federation of People with Diabetes (FFD)”.¹⁷ As expected, this required numerous revision cycles, which can be traced in the resulting translation memory (Gledhill & Zimina-Poirot, 2022; Froeliger, Gledhill & Zimina-Poirot, 2023).

As with style guides, there has been very little experimental research on position statements. Yet, given the number of available guidelines on inclusive language, it would be valuable to explore the extent to which specific inclusive language recommendations are being adopted in general discourse and by whom. In this regard, a pilot study by Duprat (2024) suggests that non-binary neologisms in French are prevalent among insiders within the non-binary community. But this kind of study is not enough: it is also necessary to examine the impact of these guidelines more objectively (e.g., by analysing language use rather than relying solely on self-reporting) and more critically: asking whether the use of inclusive language truly affects social inclusion or exclusion.

¹⁴ Available at policies.umd.edu/statement-university-values (accesses 07 May 2025).

¹⁵ Available at glucosetoujours.com/ (accessed 07 May 2025).

¹⁶ Available at commission.europa.eu/education/skills-and-qualifications/develop-your-language-skills/european-masters-translation-emt_en (accessed 07 May 2025).

¹⁷ Available at federationdesdiabetiques.org/en/who-we-are (accessed 07 May 2025).

Given that the prescriptive approach to language still plays an important symbolic role in language planning and translator training, will style guides increasingly serve only a symbolic function with the rise of generative AI assistants, which can more efficiently adapt texts to specific house styles? Or, as the example of *Language Matters* suggests, is there still room for human intervention, particularly in areas where language reformers are deliberately ‘swimming against the flow’ of mainstream discourse? It seems that while technology may streamline certain processes, human oversight will remain essential in shaping language use in areas of cultural and social significance (Froeliger, Hernández Morin, Le Floch, Look, Noonan & Volclair, 2024).

3.2. Towards a Bottom-up Approach to Defining Easy-to-Read

So far, we have examined style guides and position statements on language. However, there is clearly a significant gap between prescriptive guidelines and actual language use in practice. In this Section, we review authentic examples of documents that have been created by human authors following the Easy-to-Read (E2R) guidelines. Our aim is to identify some of the main linguistic strategies that are used in this form of simplification and to evaluate whether the observed patterns of reformulation align with the expectations of the style guide and the content of the source text. Additionally, we explore whether these strategies can be effectively replicated by generative AI or acquired as transferable drafting skills by trainee translators.

As mentioned in Maaß (2020), E2R and similar schemes, such as Easy Reading Plus¹⁸, represent a more radical form of simplification than Plain Language. The following list summarises the guidelines for E2R set out by Maaß (2020: 75):

VISUAL AND MEDIAL DESIGN: “bigger type size, sentence on new line, no word truncation, left-align”

WORD STRUCTURE: “short words, separate compounds with hyphen, no abbreviations, no passive”

VOCABULARY: “easy words, preferably no foreign words, foreign words explained”

SENTENCE STRUCTURE: “short sentences”

SEMANTICS: “no negation”

TEXT: “no lexical variation, relevant information first, use subheadings, readers addressed directly”

In some respects, the guidelines for E2R resemble those of a controlled language (Kuhn, 2014), as several rules impose formal restrictions (such as avoiding the passive voice and enforcing a limit on sentence length). However, there are significant differences as well. While controlled languages like ASD-STE100 Simplified Technical English¹⁹ typically require a correspondence between original and revised sentences, with E2R and Easy Reading Plus there is no requirement for the original segments to be represented in the

¹⁸ Available at juniorlibraryguild.com/easy-reading-plus-jep14sub (accessed 07 May 2025).

¹⁹ Available at asd-ste100.org/ (accessed 07 May 2025).

target text. Instead, these segments can be summarised or alluded to more freely, which allows for greater flexibility in the simplification process.

In the following discussion, we present a small number of examples that show the impact of these formal requirements on some of the lexical and grammatical features of E2R texts. In Sections 3.3–3.5, we analyse a sample of E2R recommendations and their impact on a specific E2R text: an appeal from the European Disability Forum (EDF) for the creation of a European disability card. We compare several versions of this appeal: a press release in formal English (European Commission, 2021), an appeal on the EDF website (EDF, 2023a), a separate pamphlet (EDF, 2023a), and the equivalent text written in Easy-to-Read (EDF, 2023b).

3.3 E2R Recommendation: Use “easy words”

The following example (1A–1B) illustrates several principles related to the E2R recommendation to use “easy words”. The authors of the E2R text have creatively interpreted this guideline by employing hypernyms (words with broader meanings, such as “countries” for “EU members”) and pantonyms (generic terms, like “people” for “political leaders”):

1A) SOURCE TEXT (European Disability Forum, 2023a [website])

We call on **European Union (EU) political leaders** to:

- Adopt an EU-wide Disability Card which ensures the **mutual recognition of disability status across Member States**, covering not only leisure, culture and sport facilities but all specific services for persons with disabilities.

1B) E2R TEXT (European Disability Forum, 2023b)²⁰

Amongst others, we ask **people who make laws and decisions in Europe** to: [...]

15. Make it possible for people with disabilities

to move freely in the European Union

This can happen by:

- **Making sure the European Disability Card**

is used by all countries in Europe

to make a difference in the lives

of people with disabilities.

²⁰ Note that for reasons of presentation we use bold to highlight specific examples. This happens to respect the guidelines for E2R, which advises against the use of capitals, colours and italics, but permits bold print for important information (Inclusion Europe, 2021: 14–17).

This example shows that lexical replacement is also simultaneously bound up with other syntactic reformulations. This includes the unpacking of pre-modified nominals into embedded clauses (e.g., “European Union (EU) political leaders” becomes “people [who make laws]”), as well as the use of light verb formulations and denominalisation (e.g., changing “ensures mutual recognition of disability status” to “making sure the European Disability Card is used by all”).

3.4 E2R Recommendation: “no passive”

The E2R recommendation to avoid the passive (“no passive”) is attractive and simple; it is encountered in many prescriptive style guides. The following example shows how a passive in the source text can be reformulated relatively straightforwardly as an active clause:

2A) SOURCE TEXT (European Commission, 2021 [Press Release])

Grant with the European Disability Card, the necessary support for persons with disabilities during the: transition periods of moving to another country to study or to work, and until **their disability has been certified by their new country of residence**.

2B) E2R TEXT (European Disability Forum, 2023b)

The European Disability Card
is a card that people with disabilities
can use when they visit or live in another country
of the European Union.
Thanks to this card
countries can recognise their disability
and give them some advantages
in culture, sports and other activities.

However, as we can see in example 1B), passive constructions are used quite regularly in E2R texts. Below are additional examples from the same brochure:

2C) E2R TEXT (European Disability Forum, 2023b)

Read a useful booklet COFACE wrote
about the rights and needs
of families of people with disabilities
This booklet **is called** ‘the SHIFT guide’ [...] Violence against women can have many forms.
For example, it may mean that women:

- **are treated** badly by their partner or their families
are forced to have sex against their will,
are threatened or feel that their lives,
 are in danger.

Prescriptive guidelines typically criticise the passive for being ‘heavy’ or for ‘obscuring agency’. But as these examples show, the construction can serve some key discourse functions. In the first example, the passive is used attributively (as a defining clause). In the second, the passive is used to topicalise “women” as the subject of several clauses, at the same time emphasising new/contrastive information, here placed at the end of each clause (“are forced against their will” etc.).

Rather than obscuring agency, the phrase “treated badly by their partner” actually emphasises this information. Therefore, we believe that the E2R recommendation for “no passive” is not always relevant. Additionally, the E2R presentation of this rule at the “Word Structure” level overlooks the fact that any active/passive reformulation affects the entire structure of the clause (and indeed, the entire paragraph).

3.5 E2R Recommendation: “no negation”

The E2R advice to use “no negation” (potentially including modality) is understandable, as multiple negations can lead to comprehension issues. The following example shows that a clause involving a double negative can be effectively reformulated as an affirmative:

3A) SOURCE TEXT (European Disability Forum, 2023a [website])

EU investments **must not finance inaccessible infrastructures**, transport or new technologies that create barriers for persons with disabilities.

3B) E2R TEXT (European Disability Forum, 2023b)

[The European Union should...]

14 **Make sure** countries in Europe

use the money they get

from the European Union

to make things better

for people with disabilities.

However, as shown in the previous examples, this strategy requires numerous simultaneous reformulations to achieve the same effect. For example, the suppression of negation involves the introduction of light verbs (e.g., *must (not)* > *make sure* and *finance* > *use the money*), and the reorientation of negative obstacles into positive causes (e.g., *create*

barriers > make things better). The use of the item ‘things’ here is also significant. This is a generic lexical item, comparable to pantonyms such as ‘people, folks, matter, stuff, events’ (rather than referring to specific persons, EU citizens etc.). The use of generic lexical items is not only indicative of a tendency to prefer hyperonyms in order to avoid specific terminology, there is also a tendency to use these words in light verb constructions, which we have seen in several of the examples above (*HAVE sex against your will*, ST *improve EU political leaders > E2R people who MAKE laws and decisions [...]*, ST *ensure recognition of disability > E2R MAKE a difference [...]*, *GIVE advantages (to people with disability) [...]*, E2R *MAKE things better*, ST *undergo violence > E2R HAVE sex against their will*, ST *must not finance > use the money they GET*, etc.).

3.6 E2R Recommendation: “short sentences”

To illustrate the E2R requirement to “use short sentences”, we have selected an example that demonstrates the replacement of one type of complexity in the source text with another in the target text:

4A) SOURCE TEXT (European Disability Forum, 2023a).

We call on the European Union (EU) leaders // to

1) Guarantee the participation of persons with disabilities in the political and public life of the EU.

4B) E2R TEXT (European Disability Forum, 2023b)

[[**What we ask**

people [[who make decisions in Europe]]]]

to do for people with disabilities

This seemingly simple paraphrase involves several interconnected reformulations. It can be observed that while the target text meets the E2R guidelines, it introduces different forms of grammatical complexity. These are characterised as: 1) grammatical density, observable in the source text as subordination and heavy pre- and post-modification, and 2) grammatical intricacy, evident in the E2R text through rank-shifted clauses, cleft clauses, parataxis, and discontinuities (especially the introductory structure: *What we ask...*). Halliday & Matthiessen (2014: 728) note that this distinction reflects the varying structural complexities in written versus oral discourse, creating an ‘oralising’ reformulation of the original text.

4. Plain Language and the Skillset for Trainee Translators

As we have seen, adapting a source text into E2R may involve several trivial transformations at the lowest level of language analysis (such as formal features and text format). However, at higher levels (including semantics and lexicogrammar), the situation is much more complex.²¹ The examples examined in Section 3 suggest that even simple transformations, like lexical replacements or changing passive to active voice, involve multiple interconnected operations that lead to significant reformulations between the source and target texts. Moreover, the segments that we looked at above represent the closest correspondences. Yet there are also many instances of reconceptualisation which obscure the relationship between the source and target texts almost entirely.

Thus, intralingual translation is clearly a complex task, whether carried out by automated assistants or language professionals. It raises several questions regarding translation teaching and research environment. While it is clearly beneficial to teach the principles of Plain Language (PL) and the guidelines for E2R in translation courses, is this truly feasible? In our previous discussions (Gledhill & Zimina-Poirot, 2022), we have found that these topics lead to further questions, particularly concerning the skills required for intra-language translation. Although language students are generally familiar with reformulation skills, especially at undergraduate level, their experience with editing and publishing can be quite limited. Additionally, translation students are often unfamiliar with concepts like PL or E2R, as these are rarely covered in traditional language courses. This raises another important question: how to address issues relating to simplification and PL in a second language with trainee translators who may not have developed these skills in their first language?

Another dimension that is barely mentioned in the literature is the relationship between simplification recommendations published in English and those provided in other languages.

Clearly the guidelines for Plain English cannot be transferred directly to other languages (as demonstrated, for example in Bouyé & Gledhill's 2024 discussion of Plain French as applied to legal texts). Furthermore, how do expectations regarding clarity and quality control procedures compare across different languages?

To this end, we examined the French version of E2R (*FALC*) provided by *Inclusion Europe*²², an organisation funded by the EU. Our initial observation is that *FALC* texts are

²¹ Our use of 'low' versus 'high' reflects a conceptual metaphor used in Systemic Functional Grammar to refer to the relative positions of the different levels or 'strata' of language (with the lower strata being seen as included in each successive higher stratum). The inherent positions of each stratum can be inferred from this definition: "Stratum. Order of symbolic abstraction along the hierarchy of stratification. The strata in the organization of language are semantics, lexicogrammar, phonology, and phonetics; context is interpreted as a stratum above language" (Matthiessen, Lam & Teruya, 2010: 205).

²² Available at inclusion-europe.eu/ (accessed 07 May 2025).

much less widespread, and those we have reviewed are unfortunately marred by fundamental language errors. Below is an example. Although this is not an official EU text, it is still surprising to encounter such a low-quality document on a website dedicated to promoting accessibility at the EU level:

*Quelle a été votre réaction lorsque
vous avez appris que vous pourriez
Pour [sic] voter lors des prochaines élections?*

Here is another example:

*Les personnes en situation de handicap
intellectuel ne sont pas différentes.
Les personnes en situation de handicap
intellectuel **être égaux** [sic] en droit.*

As far as we can determine, the practice of publishing texts without ensuring linguistic quality is quite unusual for EU-funded sites. However, such errors are revealing, especially if one considers that these texts have likely been directly ‘transposed’ from English without any effective quality control procedures in place.

5. Text Reformulation with Generative AI: Preliminary Outcomes

5.1. The Limitations of Using GAI in Reformulation

Automation serves as a means to generate summaries of complex legal and administrative documents, streamlining the process and making information more accessible. By employing advanced algorithms and natural language processing, automated tools can quickly analyse lengthy texts, extract key points, and present them in a concise format. GAI-powered assistants are also capable of generating text in “clear and simple language”. However, the capacity to generate coherent and straightforward text does not necessarily imply that automatic assistants are capable of metalinguistic reflexivity, which means understanding contextual implications, discursive functions, potential cognitive demands on readers and more. This is where human revision often becomes necessary to refine the content and ensure that it effectively conveys the intended message.

We argue that full automation applied to the reformulation of administrative and legal texts may lead to inadequate results due to several challenges inherent in this practice:

(1) Complexity of legal and administrative documents. Translating these documents into a more accessible version requires a deep understanding of the content, as well as the ability to reformulate concepts clearly and simply while maintaining legal integrity. For example, legal language often contains nuanced meanings and technical terms that are difficult for GAI to accurately interpret and simplify.

(2) Multimodal adaptation. In addition to rephrasing the text itself, Easy-to-Read legal documents often require multimodal adaptation. This involves considering features such as images, formatting, fonts, audio, and video, all of which contribute to making the content more accessible to a broader audience. It is not yet possible for GAI to efficiently combine all of these elements in order to achieve effective communication across different formats.

(3) Understanding contextual implications. While automatic assistants can simplify administrative and legal texts, they struggle to fully grasp the contextual implications of each term or concept, as well as the contextual impact of rephrasing (for example, providing examples). This lack of deeper comprehension can lead to potential mistakes in rewording, further resulting in inaccurate or incomplete interpretations of the law, unintended administrative implications, and harmful social consequences.

In summary, while automation can assist language professionals, its inherent limitations suggest that human expertise remains crucial for ensuring precision, coherence, and contextual relevance in discourse, although this balance may evolve as technology advances (Froeliger, Hernández Morin, Le Floch, Looock, Noonan & Volclair, 2024).

5.2. Finding the Balance: Human Revision vs. Automation

Given the substantial rise in automatically generated content, it is crucial to recognise that human proofreading and revision capabilities are limited. This raises the question of whether more stringent regulatory frameworks should be established to ensure that human revisions are effectively integrated into computer-assisted text generation processes. We submitted the following query to ChatGPT-4 to evaluate its automated response based on current knowledge. Here is a structured response to our prompt:²³

Should we impose stricter regulations on the production of automatically generated texts to facilitate human revisions?

²³ Our prompt (in context): *With the rise of AI assistants, there has been a significant increase in automatically generated content. Given the limitations of human proofreading and oversight, an important question arises: should we impose stricter regulations on the production of such texts to facilitate human revisions? Additionally, what strategies could be implemented to achieve this?*

Yes, imposing stricter regulations on the production of automatically generated texts could be beneficial in several ways:

Ensuring Quality and Accuracy: Regulations could mandate that all AI-generated content undergoes a human review process before publication. This would help ensure that the content is accurate, relevant, and of high quality. Humans can catch nuances, context issues, and errors that AI might miss.

Ethical Considerations: Stricter regulations could address ethical concerns by requiring transparency about the use of AI in content creation. This would help prevent misinformation and ensure that AI-generated content adheres to ethical standards.

Accountability: Regulations could establish clear guidelines for accountability. For example, they could require that content creators disclose when AI has been used, and ensure that there are mechanisms for addressing any issues that arise from such content.

How can this be done?

Establishing Guidelines: Develop and implement guidelines that require human oversight for AI-generated content. These guidelines could specify the types of content that need human review and outline the review process.

Transparency Requirements: Mandate that all AI-generated content clearly disclose its origin. This transparency can help readers understand the nature of the content and the role of AI in its creation.

Training and Standards: Provide training for individuals involved in the review process to ensure they are equipped to handle the unique challenges of working with AI-generated content. Establish standards for what constitutes adequate human revision.

Incorporating Feedback Mechanisms: Create systems for collecting feedback on AI-generated content from both users and reviewers. This feedback can be used to continuously improve both the AI systems and the human review processes.

Regulatory Bodies: Establish or designate regulatory bodies responsible for overseeing the implementation of these regulations and ensuring compliance. These bodies could also handle disputes and issues related to AI-generated content.

By taking these steps, we can better integrate human oversight into the process of creating and publishing AI-generated content, ensuring it meets high standards of quality and ethics.

We notice that this response suggests highly demanding standards for the production of texts generated by automatic assistants, focusing on clarity, accuracy and ethics. It also points out that interactions between the expert and the AI-powered assistant could include collaboration and monitoring protocols, while AI-enabled mediation processes could incorporate automatic revision tools to allow the human to review and modify auto-generated texts in compliance with institutional style guides. In parallel, in order to limit the proliferation of poor-quality texts that feed learning systems and propagate bias, it is suggested to put in place quality control and validation mechanisms.

If this perspective emerges on the basis of current semantic knowledge from a range of languages, taking into account the multilingual capabilities of ChatGPT supported by linguistic transfer (see Section 2), it appears there is a consensus among experts that they play a crucial role in maintaining a form of textual hygiene (Grigely, 1995; Cameron,

2012; Gailey, 2012), that is to say, the quality, clarity and reliability of texts generated with GAI input.

The automated response also highlights the importance of regulatory bodies in overseeing the implementation and enforcement of these regulations. Additionally, these entities are expected to manage disputes and issues related to automatically generated content. We hope that these practices will become standard in the coming years, as several suggested approaches, such as labels for machine-generated content, are already emerging (Wittenberg, Epstein, Péloquin-Skulski, Berinsky & Rand, 2024).

6. Conclusion

In this paper we have examined the potential challenges involved in applying strategies of simplification to administrative and legal content. Currently available practices range from the relatively open approach of Plain Language (PL) to more radical forms of simplification (Easy-to-Read, *FALC* and similar). Throughout this paper, we have attempted to demonstrate in concrete terms how currently available NLP tools, including those used for generating legal summaries, require not only technical proficiency but also training to operate effectively. We have further examined the interplay between generative AI, prescriptive language instruments like style guides and language policies, and the unique challenges of simplification through frameworks such as the Easy-to-Read guidelines. Our discussion of Plain Language as a competency underscores the critical interdependence of technological capabilities and linguistic expertise, while our exploration of intersubjective meta-reflexivity emphasises the evolving dynamic between AI systems and language professionals.

Ultimately, while AI-assisted translation and reformulation tools hold significant potential to revolutionise text production, the indispensable role of human experts remains clear. Their expertise is vital to ensure that these tools produce texts that are accurate, clear, and contextually appropriate, maintaining the quality and reliability required, particularly in high-stakes fields like legal communication. By fostering collaboration between technology and human expertise, we can harness the strengths of generative AI while addressing its limitations, paving the way for more effective and equitable language solutions.

In the future, one of the challenges for automation remains the question of multimodality: to get legal and administrative messages across, it is becoming increasingly useful and necessary for content to be mediated graphically/phonicallly from various technical genres into E2R or similar formats, involving changes to images, layout, font, audio, and video. Despite current technical advances, audiovisual adaptation remains basic, but end-users will probably come to expect a higher degree of remediation in the

future. For professional translators and translator trainers, this means developing new skills and integrating them into translation workflows and training programmes. For now, as argued throughout this paper, there still remains a need to address written text, as language policy is still primarily shaped by written norms. Legal content, in particular, has limits on how much it can be simplified (see Figure 2). And we have seen in Sections 2–3 that the challenges of reformulating technical legal documents into Plain Language are the same as those involved in interlingual translation. Thus, for any given text there exist many different reformulation strategies, with each modification affecting the meaning potential of the original text. With E2R writing, the content and overall message are often simplified so drastically that corresponding segments in the original text can be difficult to identify. Nonetheless, E2R texts capture the core essence of the message conveyed in more traditional prose texts.

However, it is also noteworthy that radically reformulated texts introduce new complexities, as noted by Halliday and Matthiessen (2014: 728): “written [discourse] is more complex in terms of lexical density, while spoken [discourse] is more complex in terms of grammatical intricacy”, in other words “swapping the density of written text for the intricacy of oral discourse”. This observation should not diminish our assessment of these texts; rather, it underscores their quality. As mentioned earlier, producing an E2R text is not a trivial task; it requires time, skill, and significant drafting expertise. For these texts, including those intended for individuals with cognitive impairments, human translators must generate, compare, assess, and validate various alternative formulations, all of which can indeed be suggested by automatic assistants.

In the immediate present, however, there are more practical concerns. Unfortunately, as we have noted, access to E2R texts on EU websites is almost exclusively in English, and the quality of *FALC* texts in French is often poor. How can we guarantee the editorial quality of these texts if we rely solely on source texts in English? In our previous studies, we have suggested the adoption of ‘normalisation’ within a standardised translation quality revision cycle (Gledhill & Zimina 2019; Gledhill & Zimina-Poirot, 2022), arguing that the application of a style guide / language policy should be as well-known a phase in the translation cycle as ‘post-edition’ and ‘revision’. Additionally, the predominance of high-quality E2R texts at the European level published in English raises significant issues regarding users’ rights to access quality information in their native languages. We hope that this situation will improve in the long term.

References

- Adam, Jean-Michel (2001). Types de textes ou genres de discours ? Comment classer les textes qui disent de et comment faire ? *Langages*, 35(141), 10–27. DOI: [10.3406/lgge.2001.872](https://doi.org/10.3406/lgge.2001.872).
- Aitchison, Jean (1991). *Language Change: Progress or Decay?* *Cambridge Approaches to Linguistics*. Cambridge : Univ. Press.

- Baumert, Renaud; Geslin, Albane; Roussel, Stéphanie; Schott, Stéphane & Institut francophone pour la justice et la démocratie – Institut Louis Joinet (Eds.) (2021). *Colloques & Essais. Langues et Langages Juridiques: Traduction et Traductologie, Didactique et Pédagogie*. Institut francophone pour la justice et la démocratie.
- Bock, Bettina (2018). Die Menschenrechte in ‚Leichter Sprache‘. Vergleich der Modifikationsprinzipien von zwei Fassungen. *Popularization and Knowledge Mediation in the Law/Popularisierung Und Wissensvermittlung Im Recht*. Wien, LIT, 169–200.
- Bouyé, Manon (2022). *Le Style Clair en Droit : Étude Comparative du Discours Juridique en Anglais et en Français, Avant et Après Simplification. Le Plain Language Dans la Communication Juridique Avec le Grand Public*. PhD Thesis, Université Paris Cité : Paris.
- Bouyé, Manon & Gledhill, Christopher (2024). The phraseology of legal French and legal popularisation in France and Canada: A corpus-assisted analysis. *Languages*, 9(3), 107. DOI: [10.3390/languages9030107](https://doi.org/10.3390/languages9030107).
- Cameron, Deborah (2012). *Verbal Hygiene*. Routledge. DOI: [10.4324/9780203123898](https://doi.org/10.4324/9780203123898).
- Chang, Tyler A. & Bergen, Benjamin K. (2024). Language model behavior: A comprehensive survey. *Computational Linguistics*, 50(1), 293–350. DOI: [10.1162/coli_a_00492](https://doi.org/10.1162/coli_a_00492).
- Crossley, Scott A.; Louwerse, Max M.; McCarthy, Philip M. & McNamara, Danielle S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, 91(1), 15–30. DOI: [10.1111/j.1540-4781.2007.00507.x](https://doi.org/10.1111/j.1540-4781.2007.00507.x).
- Duprat, Noah (2024). *La Non-Binarité Dans la Langue Française*. University Paris Cité : Paris.
- Edler, Cordula (2020). *Handbook for Peer-Researchers, Researchers, and Developers. Easy Reading*.
- European Commission (2021). *Information for All: European Standards for Making Information Easy to Read And Understand*. Available at inclusion-europe.eu/easy-to-read-standards-guidelines/ (accessed 07 May 2025).
- European Commission, Directorate General for Translation (2012). *How to Write Clearly*. Publications Office. DOI: [10.2782/54318](https://doi.org/10.2782/54318).
- European Disability Forum. *EDF Manifesto on the European Elections 2024: “Building an Inclusive Future for Persons with Disabilities in the EU”*. Available at edf-feph.org/publications/eppd-manifesto-2023 (accessed 07 May 2025).
- European Disability Forum. *What We Ask People Who Make Decisions in Europe to Do for People With Disabilities. Our Manifesto on the European Elections. Easy-to-Read Version*. Available at edf-feph.org/content/uploads/2000/05/EDF-Manifesto-on-the-European-Elections-2024-Accessible-Word.docx (accessed 07 May 2025).
- Froeliger, Nicolas; Gledhill, Christopher & Zimina-Poirot, Maria (2023). Intégrer des plateformes de traduction automatique neuronale dans l’enseignement de la traduction spécialisée. Froeliger, Larssonneure & Sofo (Eds.), *Human Translation and Natural Language Processing Towards a New Consensus?* (p. 19011). Fondazione Università Ca’ Foscari. DOI: [10.30687/978-88-6969-762-3/009](https://doi.org/10.30687/978-88-6969-762-3/009).
- Froeliger, Nicolas; Morin-Hernández, Katell; Le Floch, Vasilica; Looock, Rudy; Noonan, Will & Volclair, Alain (2024). «Non, l’intelligence artificielle ne remplacera pas les traducteurs et les traductrices !». *Le Monde*. Available at lemonde.fr/idees/article/2024/09/09/non-l-intelligence-artificielle-ne-remplacera-pas-les-traducteurs-et-les-traductrices_6308656_3232.html (accessed 07 May 2025).
- Gailey, Amanda (2012). Cold war legacies in digital editing. *Textual Cultures: Texts, Contexts, Interpretation*, 7, 17–5. Available at <https://api.semanticscholar.org/CorpusID:144855450> (accessed 07 May 2025).
- Gangloff, Doriane (2015). Le langage facile à lire et à comprendre. *Revue Suisse De Pédagogie Spécialisée*, 4, 44–50.
- Gledhill, Christopher; Martikainen, Hanna; Mestivier, Alexandra & Zimina, Maria (2019). Towards a linguistic definition of ‘simplified medical English’: Applying textometric analysis to cochrane medical abstracts and their plain language versions. *LCM – La Collana / The Series*, 9788879169196, 91–114. DOI: [10.7359/919-2019-gled](https://doi.org/10.7359/919-2019-gled).
- Gledhill, Christopher & Zimina-Poirot, Maria (2019). The impact of machine translation on a masters course in web translation: From disrupted practice to a qualitative translation/revision workflow. *Translating and the Computer*, 41.
- Gledhill, Christopher & Zimina-Poirot, Maria (2022). Human-machine interaction: How to integrate plain language rules in the revision cycles of neural machine translation output. *De Europa: European and Global Studies Journal*, Special Issue – 2022(1), 149–172. Available at u-paris.hal.science/hal-04050523 (accessed 07 May 2025).

- Grigely, Joseph (1995). *Textuality: Art, Theory, and Textual Criticism*. University of Michigan Press.
- Halliday; Kirkwood, Michael Alexander & Matthiessen, Christian M. I. M. (2014). *An Introduction to Functional Grammar* (4th ed.). Arnold Oxford University Press.
- Hansen-Schirra, Silvia; Bisang, Walter; Nagels, Arne; Gutermuth, Silke; Fuchs, Julia; Borghardt, Liv; Deilen, Silvan; Gros, Anne-Kathrin; Schiffl, Laura & Sommer, Johanna (2020). Intralingual translation into easy language—or how to reduce cognitive processing costs. Hansen-Schirra & Maaß (Eds.), *Easy Language Research: Text and User Perspectives* (pp. 197–225). Berlin: Frank & Timme.
- Hansen-Schirra, Silvia & Maaß, Christiane (Eds.) (2020). *Easy – Plain – Accessible. Easy Language Research: Text and User Perspectives*. Frank & Timme.
- Helm, Paula; Bella, Gábor; Koch, Gertraud & Giunchiglia, Fausto (2024). Diversity and language technology: How language modeling bias causes epistemic injustice. *Ethics and Information Technology*, 26(1), 8. DOI: [10.1007/s10676-023-09742-6](https://doi.org/10.1007/s10676-023-09742-6).
- Inclusion Europe (2021). *Inclusion Europe. Annual Report 2021*. Available at inclusion-europe.eu/annual-report-2021 (accessed 07 May 2025).
- Kantharaja, K. P.; Srinivasalu, M.; Rajashekara, M. N.; Vishvanath, Yankappa & Sanjeev, P. (2024). Exploring the impact of artificial intelligence on the future of English language editing and proofreading. *Educational Administration: Theory and Practice*, 30(1), 669–681. DOI: [10.53555/kuey.v30i1.5226](https://doi.org/10.53555/kuey.v30i1.5226).
- Kilichev, Bayramali; Abbasova, Nargiza; Gafurova, Nodira; Dadabayeva, Shirinkhon; Nurmatova, Muniskhon & Fazildinova, Sevara (2024). A development of personalized content creation technology model using NLP and AI integrated system. *4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 595–600). DOI: [10.1109/ICACITE60783.2024.10617129](https://doi.org/10.1109/ICACITE60783.2024.10617129).
- Kline, Emily A. (2023). Stolen voices: A linguistic approach to understanding implicit gender bias in the legal profession. *UCLA Women's Law Journal*, 30(1). DOI: [10.5070/L330161545](https://doi.org/10.5070/L330161545).
- Krieg-Planque, Alice (2020). Quand la communication publique travaille son expression: Les administrations à la recherche d'un «langage clair». *Politiques De Communication*, 14(1), 3–34. DOI: [10.3917/pdc.014.0003](https://doi.org/10.3917/pdc.014.0003).
- Kshetri, Nir (2024). Economics of artificial intelligence governance. *Computer*, 57(4), 113–118. DOI: [10.1109/MC.2024.3357951](https://doi.org/10.1109/MC.2024.3357951).
- Kuhn, Tobias (2014). A survey and classification of controlled natural languages. *Computational Linguistics*, 40(1), 121–170. DOI: [10.1162/COLI_a_00168](https://doi.org/10.1162/COLI_a_00168).
- Linell, Per (1998). Discourse across boundaries: On recontextualizations and the blending of voices in Professional discourse. *Text – Interdisciplinary Journal for the Study of Discourse*, 18(2). DOI: [10.1515/text.1.1998.18.2.143](https://doi.org/10.1515/text.1.1998.18.2.143).
- Lionbridge (2023). *The Lionbridge 2023 Machine Translation Report*. Available at lionbridge.com/whitepaper/the-lionbridge-2023-machine-translation-report (accessed 07 May 2025).
- Liu, Jiaxi (2024). ChatGPT: Perspectives from human–computer interaction and psychology. *Frontiers in Artificial Intelligence*, 7, 1418869. DOI: [10.3389/frai.2024.1418869](https://doi.org/10.3389/frai.2024.1418869).
- Lo Bianco, Joseph (2010). The importance of language policies and multilingualism for cultural diversity. *International Social Science Journal*, 61(199), 37–67. DOI: [10.1111/j.1468-2451.2010.01747.x](https://doi.org/10.1111/j.1468-2451.2010.01747.x).
- Maaß, Christiane (2020). *Easy Language – Plain Language – Easy Language Plus: Balancing Comprehensibility and Acceptability*. Frank & Timme. Available at library.oapen.org/handle/20.500.12657/42089 (accessed 07 May 2025).
- Mamac, Monaliza Hernandez (2024). Linguistic complexity of public legal information texts for young persons. *Text & Talk*, 44(5), 623–648. DOI: [10.1515/text-2021-0187](https://doi.org/10.1515/text-2021-0187).
- Martínez, Gonzalo; Watson, Lauren; Reviriego, Pedro; Hernández, José Alberto; Juárez, Marc & Sarkar, Rik (2023). Towards Understanding the Interplay of Generative Artificial Intelligence and the Internet. *arXiv*.
- Matthiessen, Christian M. I. M.; Lam, Marvin & Teruya, Kazuhiro (2010). *Key Terms in Systemic Functional Linguistics. Key Terms Series*. Continuum. DOI: [10.5040/9781474211888](https://doi.org/10.5040/9781474211888).
- Morrow, Garrett; Swire-Thompson, Briony; Polny, Jessica Montgomery; Kopec, Matthew & Wihbey, John P. (2022). The emerging science of content labeling: Contextualizing social media content moderation. *Journal of the Association for Information Science and Technology*, 73(10), 1365–1386. DOI: [10.1002/asi.24637](https://doi.org/10.1002/asi.24637).

- O'Brien, Sharon (2024). Human-centered augmented translation: Against antagonistic dualisms. *Perspectives*, 32(3), 391–406. DOI: [10.1080/0907676X.2023.2247423](https://doi.org/10.1080/0907676X.2023.2247423).
- Pinker, Steven (1995). *The Language Instinct* (1st Ed.). HarperPerennial.
- Pinker, Steven (2015). *The Sense of Style: The Thinking Person's Guide to Writing in the 21st Century*. Penguin Books.
- Preite, Chiara (2018). Stratégies dialogiques et transmission du savoir juridique dans le site du ministère de la justice française. *Popularization and Knowledge Mediation in the Law: = Popularisierung Und Wissensvermittlung Im Recht*, 9, 149–167.
- Rock, Frances (2007). *Communicating Rights: The Language of Arrest and Detention*. Palgrave Macmillan.
- Roffo, Giorgio (2024). Exploring advanced large language models with LLMsuite. Advance online publication. *ArXiv*. DOI: [10.48550/ARXIV.2407.12036](https://doi.org/10.48550/ARXIV.2407.12036).
- Speight, Jane; Skinner, T. C.; Dunning, T.; Black T.; Black, T.; Kilov, G.; Lee, C.; Scibilia, R. & Johnson, G. (2021). Our language matters: Improving communication with and about people with diabetes. A position statement by Diabetes Australia. *Diabetes Research and Clinical Practice*, 173, 108655. DOI: [10.1016/j.diabres.2021.108655](https://doi.org/10.1016/j.diabres.2021.108655).
- Svoboda, Tomáš (2013). Translation manuals and drafting style guides at the European Commission. *Le Bulletin Du CRATIL Centre De Recherche De L'isit*, 10, 80–106.
- Turnbull, Judith Anne (2018). Communicating and recontextualising legal advice online in English. *CHE*. Available at iris.unimore.it/handle/11380/1201050 (accessed 07 May 2025).
- van Leeuwen, Theo (1993). *Language and Representation: The Recontextualisation of Participants, Activities and Reactions*, PhD Thesis. University of Sydney: Sydney. Available at hdl.handle.net/2123/1615 (accessed 07 May 2025).
- Wittenberg, Chloe; Epstein, Ziv; Peloquin-Skulski, Gabrielle; Berinsky, Adam J. & Rand, David (2024). *Labeling AI-Generated Media Online*. Available at osf.io/b238p (accessed 07 May 2025).
- Zhang, Haochen; Dong, Yuyang; Xiao, Chuan & Oyamada, Masafumi (2024). Large Language Models as Data Preprocessors. *arXiv*.
- Zimina-Poirot, Maria (forthcoming). Computer-assisted language mediation in teaching human-centered augmented translation. *Teaching Translation in the Age of Generative AI: New Paradigm, New Learning*. Language Science Press.
- Zimina-Poirot, Maria & Gledhill, Christopher (2023). The evolutive nature of translation and teaching workflow: Textometric analysis of multiple revision cycles in a simulated website translation project. *MediAzioni*, A116-A146. DOI: [10.6092/ISSN.1974-4382/18790](https://doi.org/10.6092/ISSN.1974-4382/18790).

Note: JLL and its contents are Open Access publications under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).



Copyright remains with the authors. You are free to share and adapt for any purpose if you give appropriate credit, include a link to the license, and indicate if changes were made.

Publishing Open Access is free, supports a greater global exchange of knowledge and improves your visibility.