LONG PAPER

A study of easy-to-read adaptations in climate communication: a corpus analysis

Marina Pujadas-Farreras[1] · Sarah Anne McDonagh[1]

Abstract

Access to clear and scientifically verified information is crucial to understand the scale of the challenge of the climate crisis, allowing policymakers, governments, and individuals to better respond to it. However, for some, including people with cognitive disabilities as well as individuals who have difficulty reading, climate science can be difficult to understand with complex sentence structures and jargon obscuring important environmental information. The aim of this paper is to assess the main linguistic barriers to information about the climate crisis for people with cognitive disabilities and reading difficulties. Using the examples of the Greater London Authority's Environment Strategy Executive Report and the Northern Ireland Executive Discussion Document on a Northern Ireland Climate Change Bill, and their adaptations into Easy-to-Read as our corpus, we will examine the different morphosyntactic and lexical elements of each text to determine their overall level of complexity, such as sentence structure or vocabulary. Our results show that both Easy-to-Read versions were generally easier to understand than the standard versions. However, the Northern Ireland Easy-to-Read text was more syntactically complex with longer and more complicated sentence structures, owing to its function as a consultation document. We conclude that the main access barriers in the Northern Ireland Easy-to-Read text are related to its morphosyntactic elements, a conclusion supported by our manual analysis of the standard text. More generally, the lexical aspects of both Easy-to-Read texts showed positive results, which suggests that they may be more easily simplified compared to morphosyntactic aspects. However, more systemic and contrastive research is required to confirm our findings, in addition to user testing with people with cognitive disabilities and individuals who have difficulty reading.

Keywords Easy-to-read · Corpus analysis · Climate change · Accessibility

1 Introduction

Access to clear and scientifically verified information is crucial to understand the scale of the current climate crisis. However, climate science can be complex and difficult to understand. Moreover, how the environmental crisis is framed, explained and understood—as a fringe political issue or a looming disaster, or a variety of interpretations in between—can further impede accurate assessments of the crisis and its implications. While framing the climate crisis as solely a political issue robs it of its urgency, presenting it as a looming apocalypse can also lead to it being perceived as an inevitable catastrophe, causing some people to disengage. It is within this context that in 2022, Google announced its collaboration with the United Nations (UN) to provide scientifically verified information panels and visuals on the causes and effects of climate change (Fig. 1). According to the UN press release:

> Now, when users search for "climate change" they will find authoritative information from the United Nations in 12 languages. In addition to organic search results, Google is surfacing short and easy-to-understand information panels and visuals on the causes and effects of climate change, as well as individual actions that people can take to help tackle the climate crisis [1].

✉ Marina Pujadas-Farreras
Marina.Pujadas@uab.cat

Sarah Anne McDonagh
sarahanne.mcdonagh@uab.cat

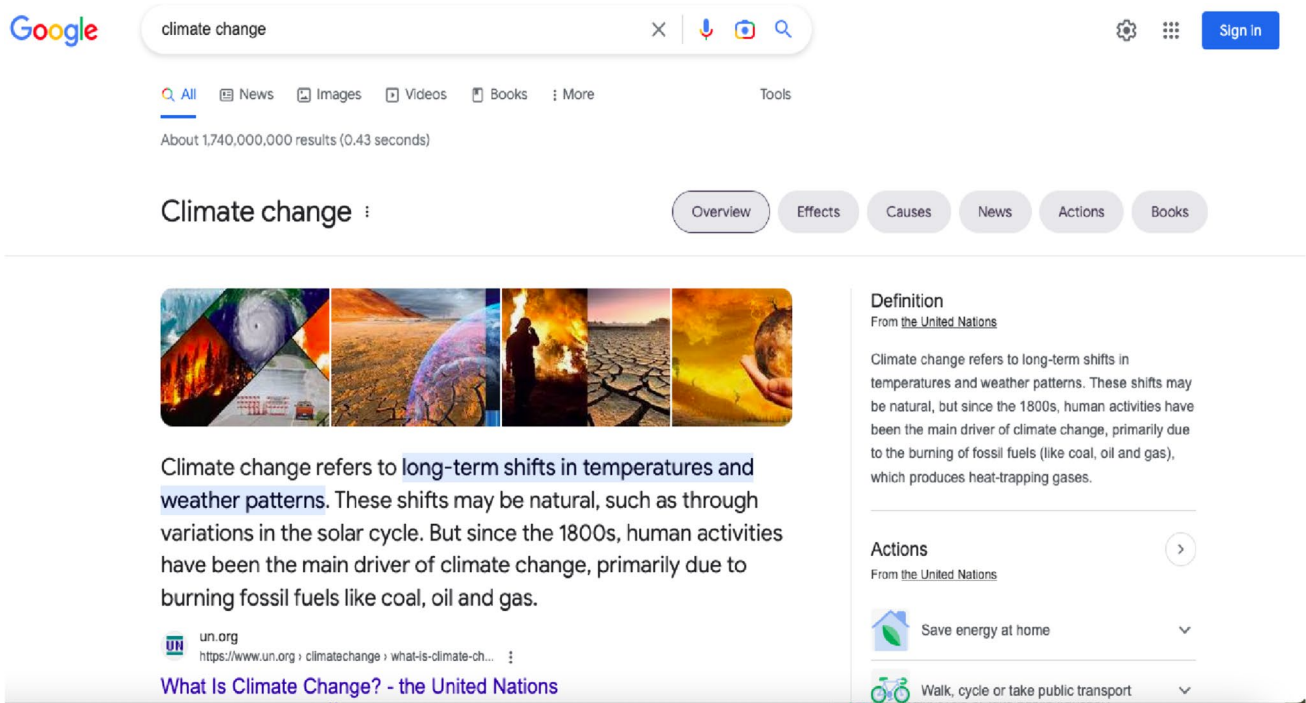[1] Universitat Autònoma de Barcelona, Barcelona, Spain

Fig. 1 Screenshot of Google panel on climate change. Source: Google (2022)

As illustrated in Fig. 2, each action is accompanied by a short explanation and supporting visuals, informing readers about the effects and causes of climate change as well as the steps they can take to tackle the crisis through

Fig. 2 Screenshot of Google panel on actions people can take to tackle climate change. Source: Google (2022)

individual and collective actions. While Google's use of Plain Language (PL), a language variation that falls under the umbrella term of Easy-to-Understand (E2U), makes the information accessible to the general non-expert reader, people with cognitive disabilities or individuals who have difficulty understanding standard written language, might still have difficulty parsing the meaning of these panels. For people with this background, Easy-to-Read texts (ER) may be more suitable. ER is a simplified language variety that aims to make a text easier to understand for people with comprehension difficulties [2]. Ensuring environmental information is accessible to a wide range of people, including individuals who have difficulty understanding standard written language, can help correctly identify, prevent, prepare for, and respond to the climate crisis in an inclusive and ultimately more effective way.

A recent trend towards clear and understandable environmental information and more disability-inclusive approaches to climate adaptation planning has emerged in environmental and sustainability policy [3, 4]. This paper will assess the main linguistic barriers to climate information specifically for people with cognitive disabilities within the context of the UK, using the examples of the Greater London Authority's Environment Strategy Executive Report and the Northern Ireland Executive Discussion Document on a Northern Ireland Climate Change Bill in Standard and Easy-to-Read (ER) [5–8]. For the purposes of this paper, we use the term "people with cognitive disabilities" to encompass a broad spectrum of potential users that include, but are not limited to, people with intellectual disabilities, Down syndrome, autism spectrum conditions, brain injuries, stroke, and age-related cognitive conditions, such as dementia. Given the diversity of this target group, providing a tailored approach to text modification is essential. Through corpus analysis, we assess the main lexical and syntactic elements of each text to determine their level of complexity. More specifically, we compare and contrast the ER texts with their standard versions to answer the following question. Are the ER texts, in fact, easier to understand? For the purposes of our analysis, we refer to the London Environmental Strategy Executive Summary as "London ST" and its ER version as "London ER". Similarly, we use the shorthand "N. Ireland ST" to refer to the Discussion Document on a Northern Ireland Climate Change Bill and "N. Ireland ER" for its ER version. Finally, we conclude this paper by examining the case for the adoption of ER conventions in climate communication, discuss the inherent tensions between comprehensibility and the complexity of climate science, and propose future avenues of research.

2 Background to this study

2.1 Access to climate information for people with disabilities

Despite the disproportionate effects of climate change on the lives and well-being of people with disabilities, they are largely ignored in climate crisis mitigation and adaptation planning [9]. This is a failure on the part of governments to fulfil their obligations to respect and uphold the rights of persons with disabilities, as agreed under the terms of the Paris Agreement and the Convention of the Rights of Persons with Disabilities (CRPD) [10, 11]. According to a recent report by Jodoin et al., only 45 out of 192 signatories of the Paris Agreement explicitly mention people with disabilities in their national climate adaptation policies [9]. Where policy documents do mention people with disabilities, it is normally cursory in nature and are lacking in concrete proposals. Without concrete measures in place to include people with disabilities in climate planning, they stand to be further disproportionately negatively affected by the ongoing climate crisis [12]. Increased exposure to climate extremes and its effects, coupled with a lack of preparedness measures, places people with disabilities at a significant disadvantage to their able-bodied counterparts, exacerbating already existing social inequalities [12]. In situations of environmental disasters, people with disabilities face barriers, such as inaccessible resources and emergency housing, information and services, all of which put their lives at risk [13, 14].

All of this is exacerbated by the fact that people with disabilities often lack access to basic information. Inaccessible government websites, lack of sign language interpreters at press conferences, inaccessible visuals (i.e., graphs and images), and complex jargon all render vital climate information inaccessible to people with disabilities. To begin to address this, the provision of accessible environmental information in a variety of different formats, including ER, can improve the ability of people with disabilities to adapt and respond to climate-induced extreme weather events, as well as participate in climate planning, adaptation and activism at local, national and international levels. For people with cognitive disabilities, accessible information can be delivered through various creative and engaging ways, such as ER summaries and visual aids. These formats can enhance understanding of climate change and its impacts, empowering individuals with cognitive disabilities to make informed decisions, adapt to shifting weather patterns and participate in environmental preservation.

2.2 The potential of easy-to-read in environmental communication

As an extension of "Easy to Understand", ER is a form of accessible communication that seeks to enhance comprehensibility and perceptibility of any written text [15]. In line with the ISO standard, ER can be defined as a "language variety in which a set of recommendation regarding wording, structure, design and evaluation are applied to make information accessible to persons with reading comprehension difficulties for any reason" [15]. ER is primarily designed to cater to the needs of individuals with cognitive disabilities. Due to the diversity of this group, a tailored approach to text modification is essential to ensure effective comprehension and engagement for a diverse range of readers [16]. For example, a person with dyslexia might find it difficult to decode certain words or phrases within a text, whereas someone with a developmental disorder may not fully understand the overall meaning of a text [12]. Consequently, the challenges of understanding and reading a text vary according to the individual needs of the reader. Although ER materials are primarily aimed at people with cognitive disabilities, their use extends beyond this group to include people across different age ranges with and without disabilities. This includes adults and children who have difficulty understanding standard written language, as well as non-native speakers seeking to enhance their language proficiency [17]. By reducing linguistic complexity and using visual aids, ER enhances comprehension and facilitates easer navigation of the written text for users.

As an instrument of inclusion, ER shares many attributes with PL, a language variation primarily aimed at non-expert readers [18]. However, unlike PL, which mainly caters to people with a standard adult reading level, ER is specifically tailored to people who have difficulty reading and understanding standard language. The certification of a document or text as ER is governed by strict conventions related to its content and form [2, 15]. These include the use of short sentences, simple and direct language, explanation of difficult terms or phrases, and summaries of key information, large text sizes, clear sections and headings, with sentences presented in the form of bullet points and usually accompanied by supporting images. The paratextual elements of ER, such as illustrations and images, also follow specific rules related to their layout and function within a text [18]. According to guidelines [2, 15], for a text to be considered ER, it must firstly undergo a verification process completed by ER users from user organisations to check its comprehensibility.

While the adoption of ER has steadily increased internationally [19], its application remains uneven, with inconsistencies in which ER rules are applied across different countries and jurisdictions, often owing to linguistic differences, such as language-specific rules, and inconsistencies between different ER guidelines and validation practices [19–21]. Although there is abundant research within academia, such as [22–24], this research may not always reach practitioners due to a lack of awareness or understanding of the publication language. Furthermore, there are aspects in ER research that are still lacking, such as the use of images [20]. As pointed out by González-Sordé and Matamala, "without evidence to base the creation of guidelines on, no hierarchy of recommendations has yet been set, and a few inconsistencies between different publications can be found" [20].

In the context of the UK, the focus of this article, ER practices have steadily increased in response to the growing disability rights movement, which challenged the historical marginalisation of disabled people and instead advocated for their inclusion in society. Guided by the social model of disability [25], which posits that society's barriers, not personal impairment, create disability, campaigners called for the removal of these barriers and increased opportunities for disabled people to fully engage in society [...]. As suggested by Chinn and Buell, equitable access to texts and information forms a key tenet of this goal, enabling disabled individuals to participate fully in all aspects of civic life, including climate planning, mitigation and action [26].

The following section examines the application of ER to two policy documents taken from the Department of Agriculture, Environment and Rural Affairs within Northern Ireland Executive and the London Mayor's Office, part of the Greater London Authority, both of which focus on climate and environmental issues.

3 Methodology

For the purposes of this study, we conducted a corpus analysis of two government documents, both of which were adapted into ER. We chose corpus analysis as our preferred method to detect patterns and trends in the ER and standard versions of both texts. By using different metrics, corpus analysis provided us with the opportunity to evaluate the comprehensibility of the standard and ER versions of each text. Although corpus analysis has many benefits, such as providing valuable comparative data, it also has limitations, such as an overreliance on quantitative data and lack of nuance or context [27]. These limitations were partially addressed by conducting a manual analysis of both documents and their adaptation into ER to either contextualise the findings. Our corpus analysis was based on the methodology followed by Arias-Badia and Matamala in their EL corpus analysis [28], in which we analysed the morphosyntactic and lexical elements of each text, which is explained in more detail below.

These documents were selected for analysis because they both focus on the topic of sustainability. While both

documents broadly focus on the topic of sustainability, they have different aims. The London text is a policy document that outlines the Mayor of London's environmental strategy for the city. It details various approaches to address issues such as air quality, waste, energy use, and noise pollution. Additionally, it covers policies that promote green infrastructure, climate change adaptation and mitigation, and a low carbon economy. In contrast, the N. Ireland text is a consultation document that provides the rationale for the introduction of legislation on climate change specific to Northern Ireland. It outlines the current national and international legal context, before moving on to set out possible legislative options to address climate change in the region. Unlike the London text, which is primarily informative, the N. Ireland text encourages the reader to participate in the consultation process. As a result of its dual purpose to both inform and persuade the reader of the necessity of environmental legislation in the region, the N. Ireland text has more complex aims than the London text. Another difference between the two ER texts relates to their target audiences. According to the Northern Ireland Department of Agriculture, Environment and Rural Affairs' website, the N. Ireland ER version is intended for "younger citizens and those with reading difficulties" [29]. Conversely, the London ER version is primarily targeted at ER readers, as indicated by the ER logo on its front cover [8].

As shown in Table 1, both ER texts are significantly shorter than their corresponding standard versions. The N. Ireland ER text has a total of 1,772 tokens (i.e., total number of words), whereas the N. Ireland ST has a total of 13,748. Regarding types, which refers to the number of unique words in a text, the N. Ireland ER has 475, while the standard version has a total of 2,123. For example, the word *climate* appears 48 times in the N. Ireland ER text, but only counts as one type. As a point of comparison, the London ER text has 1,577 tokens and 493 types while its standard version has 7,227 tokens and 1,435 types, as shown in Table 1.

We used the online tool SketchEngine to analyse the text. Through this tool, we obtained an automatic lemmatisation of the text, Parts-of-Speech (PoS) tagging and frequency lists. The process of lemmatisation involves the reduction of the words in a text to their corresponding lexemes, which denotes the base form of the words [30]. For example, the

Table 1 Text information

Text	Tokens	Types	Year	Source
N. Ireland ER	1,772	475	2020	Department of Agriculture, Environment and Rural Affairs' website
N. Ireland ST	13,748	2,123	2020	
London ER	1,577	493	2018	London City Hall website
London ST	7,227	1,435	2018	

verb lemma *be* has the forms *is, are, was*, etc. Similarly, PoS tagging is a type of annotation in which each word is assigned a grammatical category (nouns, verbs, adjectives) [30].

By conducting concordance analysis, SketchEngine enabled us to differentiate between homographs. To analyse the morphosyntactic aspects of each text, we chose the following three parameters. We used SketchEngine to automatically analyse mean sentence length, mean verbs per sentence, and PoS distribution, each of which helped us determine whether the texts adhered to ER recommendations of short and simple sentences and a more direct verbal language [31, 32]. This was complemented with a manual analysis to assess how information was presented in the texts, such us use of bullet points, images, or additional explanations. For the lexical aspects, we focused on the following parameters: corpus aboutness, mean word length and lexical density (Type-Token Ratio or TTR, vocabulary richness and information load). Finally, we used two readability indices (Gunning Fog index, Flesch Reading Ease).

According to Oakes, corpus aboutness is used "(to) show to what extent the individual word types in a corpus typify the corpus as a whole", that is, to know "the set of words found to score most highly" [33]. To determine the corpus aboutness of each text, we took the 30 most frequently used words (nouns, verbs, adjectives, and adverbs) from each text using SketchEngine to assess the frequency of uncommon or complex words, such as jargon (i.e., net zero) within the texts. If difficult words were found, we analysed how this difficulty had been solved (i.e. by explaining the word). We also compared the results of the corpus aboutness of each ER text with its corresponding standard version to determine the level of correlation between the original and the ER versions. Finally, we conducted a comparative analysis and a manual analysis of the corpus aboutness of both ER and standard texts to determine whether unexplained infrequent words, such as specialised terms, have a significant presence in the text. Additionally, analysing corpus aboutness also revealed which words were frequently repeated across both the original and ER texts.

In order to study lexical variation in each text, we computed the lexical density by calculating the TTR, vocabulary richness, and information load. TTR measures lexical density by dividing the number of types with the number of tokens. A high type/token ratio indicates that the text is lexically diverse, whereas a lower result suggests that there are a lot of repetitions [30]. Vocabulary richness is measured by dividing the total lemmata by the number of tokens, and information load is then measured by dividing content words (noun, adjectives, adverbs, and verbs) by the tokens.

Finally, we used two readability indices, the Gunning Fog Index and the Flesch Reading Ease to analyse the morphosyntactic and lexical elements of each text. While these

indexes are not specifically designed for people with cognitive disabilities, they have been used in studies with this user group. For example, the Gunning Fog Index has been used in other studies to assess the difficulty of understanding content like museum AD and film AD [28, 34]. The Flesch Reading Ease index has been used in reading comprehension studies with users with cognitive disabilities [35–37].

To calculate the Gunning Fog Index of each text, we used an online calculator [38] drawing on a random sample of 100–150 words in each text, as recommended in ReadabilityFormulas [39]. As a measure of readability, the Gunning Fog index estimates the years of formal education needed to understand a text on first reading [40]. The parameters used to calculate the Gunning Fog index are the number of major punctuation marks, the total number of words, and the number of words with three syllables or more [38]. In contrast, the Flesch Reading Ease calculates a text's level of readability based on the average sentence length and the average number of syllables per word [39]. Here too, we used a random sample taken from each text to measure readability, after which we performed a manual qualitative analysis to examine how difficult words or concepts were rendered in the ER texts. We expected to find a difference between the use of verbs in the ER texts and the standard texts, following the E2U recommendations that encourage a verbal style over nominal style in ER texts [31, 32]. We also expected to find similarities in the simplification strategies of the two ER texts. Regarding the corpus aboutness, we initially expected to find a higher correlation between the ST and ER versions of each text, followed by a higher correlation between the two ER texts. A summary of this methodology is presented in Fig. 3.

4 Morphosyntactic analysis

4.1 Sentence complexity

To measure sentences complexity, we analysed both ER texts and their standard versions. Both standard texts had a mean sentence length of 28 words. As a point of comparison, the English Web 2020 text integrated into SketchEngine scores 21.4 on average [41]. Consequently, the ST texts had considerably longer sentences than the English Web 2020 average, which may be due to the ST texts' functions as policy and consultation papers. In contrast, the N. Ireland ER text had a mean sentence length of 20 words, whereas the London ER text had 11.8. Both ER texts had considerably reduced the number of words per sentence on average in comparison to the standard versions. However, even with the reduced word count, the N. Ireland ER text maintained a similar sentence length to the average English Web text. This suggests that the length of the sentences remained closer to that of standard English according to English Web 2020, which aligns more closely with PL.

The number of verbs per sentence is also a useful indicator of sentence complexity. As shown in Table 2, London ST had the highest average of the four texts, with four verbs per sentence, while the N. Ireland ST text had three verbs per sentence. Similar to the results of mean sentence length, the London ER reduced the number of verbs per sentence to an average of 1.7, which indicates simplification. This finding is further corroborated by the results of our qualitative analysis, which found that the London ER text favoured simple sentences with one piece of information per sentence, as shown in Fig. 4.

Most of the complex sentences found in the London ER had subordinate clauses that serve as the object of the sentence, some of which are difficult to avoid, particularly when using verbs that express thoughts or opinions [42]: "The Mayor wants to know what you think" or "The Mayor wants to protect against floods". In contrast, the N. Ireland ER text had 2.5 verbs per sentence on average, suggesting there was a predominance of more complex sentences.

Our qualitative analysis revealed that in addition to complex sentence structures, the N. Ireland ER text also used negatives forms, which are generally discouraged in ER guidelines [32, 42]. Although it is not feasible to eradicate

Table 2 Sentence complexity statistics

Text	Tokens	Types	Mean sentence length	Mean verb per sentence
N. Ireland ER	1,772	475	20.1	2.5
N. Ireland ST	13,748	2,123	28.0	3.1
London ER	1,577	493	11.8	1.7
London ST	7,227	1,435	28.2	4.0

Fig. 3 Summary graph of the methodology

Morphosyntactic analysis	Lexical analysis	Readability index
☐ Sentence complexity (SketchEngine)	☐ Corpus aboutness (SketchEngine)	☐ Gunning Fog Index
☐ Manual analysis	☐ Word length	☐ Flesch Reading Ease
	☐ Lexical density	

London's environment

The environment is:

● The air we breathe

● The water we drink

Fig. 4 Screenshot of London ER text, which presents one piece of information per sentence accompanied by a visual

all negatives from a text, combining them with complex sentence structures lead to sentences such as the following, which are very difficult to understand.

> The total greenhouse gas emissions in Northern Ireland must be no more than the amount of greenhouse gases that are removed from the atmosphere (in Northern Ireland), by 2050.
> It is unlikely that Northern Ireland will be able to achieve net-zero by 2050, due to the characteristics of our society, economy, and infrastructure.

Simplifying these sentences would improve their readability, even if the negative form must be maintained.

As previously mentioned in Sect. 2.1, there are varying opinions regarding the use of lists and bullet points to organise information in ER guidelines. While some guidelines

recommend their use [18, 32, 43], others caution against overusing them [42]. In publications such as Rink [44] the use of bullet points lists is mentioned as well as helpful for the readers, especially when complex sentence structures are also used. Although bullet points were used in the London ER text to separate ideas, as shown in Fig. 3, the N. Ireland ER employed them to separate paragraphs (see Fig. 5), which could be viewed as excessive.

4.2 Manual analysis

A preliminary PoS analysis indicates that both ER texts use a higher percentage of verbs than the ST versions. This suggests a more verbal style instead of a nominal style, adhering to ER recommendations [31, 32] and is further confirmed by a manual analysis.

Moreover, in both ER texts, a substantial portion of lexical verbs is employed, constituting 71% in the London ER text and 75% in the Northern Ireland ER text. Excluding auxiliary verbs, the most frequently used verbs in the London ER text include "make", "want", "use", "help" or "recycle". In contrast, the Northern Ireland ER text features recurrent use of verbs, such as "achieve", "include", "reduce", "set" or "get".

Upon examining both ER texts, distinct strategies for clarification become apparent. In the N. Ireland ER text, complex terms are clarified by incorporating a synonym in parentheses alongside the word. There are a total of seven instances of this strategy, and if a word is repeated, the synonym is also repeated. For instance, the term "legislation", followed by ('law'), recurs five times. It is unclear why the simpler synonyms are not directly used in the text, especially when dealing with terms that are not specialised and challenging to replace. Parentheses are also used to give more

Fig. 5 Use of bullet points to separate paragraphs in N. Ireland ER text

Climate change legislation (Law)

· Northern Ireland does not have its own climate change law, unlike all other parts of the UK. Northern Ireland is currently tackling climate change through a UK-wide Climate Change law, called the 'UK Climate Change Act 2008'.

· In 2019, the UK Climate Change Act 2008 was updated by the UK Government, to include the requirement that emissions of Greenhouse Gases must be reduced enough to achieve 'UK Net Zero', by the year 2050.

context. For example, while speaking of "gas heating", it is added that "gas heating causes lower greenhouse emissions". Additionally, there are three occurrences of in-text explanations, where the clarification of a concept is seamlessly integrated into the text rather than presented as an addendum.

'Climate Change Mitigation' means taking action to reduce the causes of climate change - such as reducing greenhouse gases being emitted into the atmosphere, or using 'carbon sinks' to store greenhouse gases.

However, these in-text explanations and the contextual parentheses employ terms such as "greenhouse gases" or "greenhouse emissions" and "carbon sinks", which are not explained anywhere in the text. This makes the clarifications more confusing, because they add more unknown vocabulary for the reader. It is also worth noting that different terms for the same concept are used (e.g.: "greenhouse gases" and "greenhouse emissions").

On the other hand, the London ER text adopts a different approach by including concept explanations within squares, highlighting the explained concept in bold. This particular strategy is employed in four instances (Fig. 6).

In-text explanations are also used, five times in total. For example: "The economy is how money is made and spent

other effects of climate change. This includes flooding, high temperatures and **drought**.

Drought is when it hasn't rained for a long time and plants start to die.

Fig. 6 Example of a clarification strategy in the London ER text

and the effect this has". In contrast to the N. Ireland ER text, the explanations of this text use easier vocabulary and structures, and do not use unexplained terms to elucidate new ideas. For instance, before explaining the concept "low carbon circular economy", the text first provides explanations for "economy" and "circular economy."

5 Lexical analysis

5.1 Corpus aboutness

As shown in Table 3, the most frequently used words in the text were those related to the topic of sustainability. In the case of the N. Ireland ER text, some of the most frequently used words were nouns like "climate", "change" and "gas". Besides generic verbs such as "to be", "to have" and "to do", there were also verbs such as "to reduce", "to achieve" or "to help". An n-gram search of the words featured in the corpus aboutness showed that phrases such as "climate change", "greenhouse gas" and "net zero" were also frequently used. Of these, only "net zero" was defined: "mean[ing] that the UK will limit the amount of greenhouse gases being released ('emitted'), to a level no more than the amount of greenhouse gases that are removed from the atmosphere". However, this explanation itself is problematic, as it is a complex sentence that renders the meaning unclear. Additionally, the phrase "greenhouse gas" is not explained anywhere in the text, which could make the above explanation confusing to some readers.

The corpus aboutness of the London ER text yielded different results. The most frequently used verbs in the London ER text were: "to do", "to make", "to want", "to have", which are less related to sustainability and more generic. The nouns that appeared most frequently were explained within the text, as was the case for "environment". The n-gram search revealed expressions like "climate change",

Table 3 30 most frequently used words in each text

	Corpus aboutness
N. Ireland ER	climate (48), change (47), Northern Ireland (41), be (39), law (22), gas (21), greenhouse (21), UK (17), document (16), emission (13), do (13), discussion (13), bill (12), net (11), zero (11), have (9), target (9), reduce (7), achieve (7), include (7), expert (7), set (7), atmosphere (6), local (6), option (6), make (6), get (6), help (6), advice (5), requirement (5), respond (5)
N. Ireland ST	be (310), climate (291), change (281), Northern Ireland (255), UK (148), emission (145), target (139), bill (87), GHG (72), set (71), reduction (67), carbon (61), act (57), provide (53), document (51), discussion (50), have (47), option (43), advice (42), net (40), page (39), term (39), long (38), government (37), also (37), legislation (36), reduce (35), response (35), budget (34), take (33)
London ER	be (41), London (26), make (22), mayor (20), people (19), more (17), want (13), air (12), waste (12), green (12), environment (11), electricity (11), new (11), have (10), good (9), use (8), thing (7), do (7), water (7), carbon (6), space (6), building (6), quality (5), home (5), way (5), help (5), health (5), noise (5), economy (5), transport (4)
London ST	be (165), London (155), city (57), emission (50), Londoner (50), energy (50), new (46), mayor (45), help (43), green (40), make (39), quality (32), waste (31), reduce (31), have (30), air (30), environment (30), people (27), more (27), change (26), climate (25), carbon (24), water (24), space (24), improve (24), business (23), noise (23), infrastructure (23), good (23), use (22)

Table 4 Corpus aboutness correlation between the text

	N. Ireland ER (%)	N. Ireland ST (%)	London ER (%)
N. Ireland ER			
N. Ireland ST	53.3		
London ER	16.7	10.0	
London ST	26.7	23.3	66.7

Table 5 Mean word length of each text

Text	Tokens	Syllables per word	Characters per word
N. Ireland ER	1,772	1.16	5.03
N. Ireland ST	13,748	1.48	5.06
London ER	1,577	1.14	4.92
London ST	7,227	1.69	4.75

"zero carbon" and "dirty fuel", each of which was explained in the text. For example, "zero carbon means that dirty fuels like coal and oil will not be used to make electricity or drive our transport". The difference in corpus aboutness is likely due to the different aims of each text: the London ER text primarily explains the Mayor of London's environmental policy, whereas the N. Ireland ER text argues the case for environmental legislation for the region.

A comparison of the corpus aboutness of each text revealed that, as expected, both ER versions shared most words with their ST versions. However, the N. Ireland ER text and the London ST had the second-highest corpus aboutness correlation, with 26.7% shared words. This was surprising as we expected a higher correlation between the two ER texts. In contrast, the two ER text corpus aboutness only shared 16.7% of similar words (Table 4). This indicates that the N. Ireland ER text is closer to the London ST text in terms of ease of reading, which suggests it might be best categorised as EL + or a simplified PL rather than ER .

5.2 Word length

Research suggests that words that range from 3 to 5 characters are generally easier to understand than words over 5 characters [45]. Consequently, the length of the words in each text has an impact on comprehension. In general, the ER texts used short monosyllabic words, of approximately 5 characters, which is close to the average word length in English of 5.1 characters per words [46]. Longer words in the N. Ireland ER text included "atmosphere" and "requirement" and, in the London ER text, "environment". While the London ER text explained these difficult words to the reader, this was not the case in the N. Ireland ER text, which neglected to define certain terms such as "low carbon economy" [6]. While word length is important, a reader's familiarity with words is also an important factor to consider (Table 5).

5.3 Lexical density

When interpreting the TTR results of each text, it is important to consider the number of tokens in each text, as the longer a text is, the lower the TTR [47]. While initially, it appears that the standard text has a lower TTR score than the ER versions, which have a higher lexical density, in reality, it is the contrary. If we consider the number of tokens, the N. Ireland text has a difference of only 0.11 in the TTR results, while the ER text is 12% of the complete length of the standard version. Similarly, the London text has a difference of only 0.12, while the ER text is 21% of the length of the standard version. This indicates that the text had a lot of repetitions, following ER recommendations [2, 15] of using the same words consistently throughout a text to refer to a concept or an object.

Similarly, the results of vocabulary richness are also more complicated than they initially appear. While it seems that the ER text is richer, in fact the results are higher because the total length of the text is shorter. To have a better understanding of these results, we must examine the lists of the corpus aboutness. In the N. Ireland ER text, the 30 most frequently used words account for 47% of the total words of the text (including repetitions). In contrast, the corpus aboutness of the standard version is 34.6%. Similarly, in the London ER text, the most frequently used words make up 39.2% of the text, while in the standard version it is only 27.7%. Again, this suggests that the ER texts use the same words, instead of synonyms, which is encouraged in ER guidelines [2, 42, 48].

Finally, the results show that both the ER text and the standard text have similar information load, ranging from 0.53 in both ER texts, to 0.56 and 0.61 in the standard texts, as shown in the table below. According to Johannson, most written texts have a lexical density (referred to by us as information load) of 40% or higher [49], which indicates that our corpus ST texts are average for a standard written text. There are currently no guidelines on the optimal information load for ER texts. However, in both ER versions there is a reduction in information load: 0.08 (8%) in the London ER, and 0.03 (3%) in the N. Ireland ER (Table 6).

6 Readability indexes

To evaluate the readability of the text, we used the Gunning Fog Index and the Flesch Reading Ease. According to the Gunning Fog Index, texts that score between 5 and 10 are considered ER, which is the equivalent of 5th grade up to

Table 6 Lexical density results

Text	Tokens	TTR	Vocabulary richness	Information load
N. Ireland ER	1,772	0.26	0.20	0.53
N. Ireland ST	13,748	0.15	0.10	0.56
London ER	1,577	0.31	0.23	0.53
London ST	7,227	0.19	0.14	0.61

Table 8 Flesch Reading Ease score table

90–100	Very easy
80–89	Easy
70–79	Fairly easy
60–69	Standard
50–59	Fairly difficult
30–49	Difficult
0–29	Very confusing

Table 7 Gunning Fog Index score table

17: College graduate
16: College senior
15: College junior
14: College sophomore
13: College freshman
D A N G E R L I N E
12: High school senior
11: High school junior
EASY READING BELOW THIS LINE
10: High school sophomore
09: High school freshman
08: 8th grade
07: 7th grade
06: 6th grade
05: 5th grade

Table 9 Readability indexes results

Text	Tokens	Gunning fog index	Flesch reading ease
N. Ireland ER	1,772	13.39	53.5
N. Ireland ST	13,748	22.76	11.3
London ER	1,577	4.53	84.1
London ST	7,227	12.63	56.8

each index uses to calculate the score. The Gunning Fog index considers the number of major punctuation marks, the total amount of words, and the number of words with three syllables or more [38] while the Flesch Reading Ease considers average sentence length and average number of syllables per word [39].

It is worth noting that both readability indices indicate a higher level of similarity between the N. Ireland ER and the London ST text, than between the N. Ireland ER and the London ER text. This finding aligns with our corpus aboutness analysis, providing further evidence to support our suggestion that the N. Ireland ER text may be closer to EL + or a simplified PL.

high school sophomore (U.S.) reading level or Year 6 to Year 11 (UK). Between 11 and 12 is considered standard; and, from 13 upwards, is considered difficult and corresponds to university undergraduate and graduate level reading ability [50] (Table 7).

Both ER texts are significantly easier to read when compared with the standard versions. However, while the London ER text has a readability score of 4.53 on the Gunning Fog Index, which indicates that it is very easy to read, the N. Ireland ER text has a score of 13.39, which is considered difficult, although close to standard. The Flesch Reading Ease readability scores are interpreted as follows in Table 8 [39]:

As shown in Table 9, the N. Ireland ER text shows a similar score to the Gunning Fog result; 53.5 is considered fairly difficult, although as with the previous results, it is close to the standard readability score. However, the result is lower than the London ST text, which has a score of 56.8. This result is striking because according to the Gunning Fog Index, the N. Ireland ER text is more difficult than the London ST, but according to the Flesch Reading Ease, N. Ireland ER text is easier than the London ST text.

The London ER text scored 84.1 on the Flesch Reading Ease, which is considered easy. In contrast, the Gunning Fog Index results of the London ER text scored extremely easy. The difference in the score is explained by the parameters

7 Summary of findings

The results of our corpus analysis are summarised in Table 10.

The results of the automatic morphosyntactic analysis revealed that both ER texts have shorter sentences than the ST texts. Nevertheless, the N. Ireland ER text sentence length was considerably longer than that of the London ER text, averaging 20.1 words per sentence, which is closer to PL. The N. Ireland ER text also used substantially complex sentences, with an average of 2.5 verbs per sentence, compared to the 1.7 of the London ER text. Furthermore, the N. Ireland ER text paired complex sentence structures with negative structures, which further complicates the text. Both ER texts also used bullet lists, however they used them differently. While the London ER texts used bullet points to break down information, as per ER recommendations [18,

Table 10 Summary table of the analysis

	N. Ireland ER	London ER
Short sentences	X	✓
Simple sentences	X	✓
Bullet lists	X	✓
Verbal style	✓	✓
Common vocabulary	✓	✓
Explanation of difficult words	X	✓
Short words	✓	✓
Low vocabulary richness	✓	✓
Readability	X	✓

32, 43], the N. Ireland ER used them to separate different paragraphs, which may be considered an overuse. The preliminary PoS analysis indicated that both ER texts used a more verbal style, as per recommendations [31, 32], and the manual analysis showed different clarification strategies: in-text explanations, and synonyms and clarifications in parentheses in the N. Ireland ER text; and in-text explanations and explanations in boxes in the London ER text. The analysis revealed that the clarification strategies in the N. Ireland ER text were insufficient, further proving its closeness to PL rather than ER. On the other hand, the London ER text employed the clarification strategies successfully, ensuring an easy-to-understand text.

The results of the lexical analysis showed that both ER texts used less specialised vocabulary compared to the standard versions, indicating that the ER texts are easier than the standard versions. However, the N. Ireland ER text failed to clarify the meaning of some of the more complex terms it used. Finally, the readability scores indicated that the London ER text can be considered ER, whereas the N. Ireland ER text may pose a challenge for some readers.

Owing to this complexity, we concluded that the N. Ireland ER text is closer to PL. This suggestion is further supported by the similarities between N. Ireland ER text and London ST. The higher level of difficulty in the N. Ireland ER compared to the London ER may be the result of failing to validate the ER text with target users, in addition to a lack of understanding of ER guidelines and requirements. On the other hand, the London ER text shows all the characteristics of a correctly done ER text. One possible reason for the higher level of difficulty with the N. Ireland ER compared with the London ER may relate to their different target audiences. While the London ER text is explicitly aimed at people with cognitive disabilities, the N. Ireland ER text is aimed at a broader target group, specifically, "younger citizens and those with reading difficulties" [29]. Consequently, the N. Ireland is not intended for the traditional ER target group. Based on our analysis, we conclude that the main access barriers to comprehension were related to

morphosyntactic aspects, which is supported by the N. Ireland ER text negative results and further reinforced by our analysis of the ST text. While lexical aspects should also be simplified, our analysis of both ER texts demonstrated positive results more generally, which suggests that the lexical elements of a text are easier to simplify compared to the morphosyntactic.

8 Limitations

One limitation of our corpus analysis is that we used a small sample of two standard texts and their respective ER versions. Furthermore, each text has a different communicative aim: the London text is primarily informative, whereas the N. Ireland text is discursive. To gain a more thorough understanding of ER usage for environmental communication, future research should include a more extensive corpus of texts on sustainability with varying communicative aims and comparable text types. This would enable a more nuanced analysis of the different approaches to discussing and conveying environmental issues through ER, targeting diverse audiences.

9 Conclusion

While the primary focus of this article is on individuals with cognitive disabilities, issues that relate to information access are systemic in nature and extend beyond this particular demographic or field of communication. The European Union, for example, has emphasised that the availability of clear communication continues to be a persistent problem, even within its own organisations and bodies [51]. However, simplification, while at times essential, may result in the loss of the necessary complexity and subtleties needed for discussions about the environmental crisis. As a result, there exists a fundamental tension between simplification and the detail necessary in communicating complex ideas and fields, such as climate science. However, it should be noted that ER is just one method to simplify communication. As highlighted by González-Sordé and Matamala "since Easy Language can serve everyone, everyone should be able to choose between information in Easy Language and standard language" [20].

Consequently, we are not advocating for the replacement of standard texts with ER. Instead, we propose ER as an option for those who require it. By adopting a disability-informed approach to climate communication, we can educate people about the environmental crisis and related health risks through simple, clear, and easy to understand messaging that is available in a range of different formats. More specifically, improving the effectiveness of climate

communication for people with cognitive and intellectual disabilities as well as for individuals who have difficulty reading requires a tailored approach that draws on ER conventions and best practice. Nevertheless, as the results of our analysis show despite the prevalence of ER guidelines, deviations occur between supposed ER texts due to their different functions and variability in current ER guidelines. Additionally, more empirical evidence is required to either validate or refute specific ER guidelines, specifically related to questions of lexical density or information load to ascertain the desired average for ER texts [49]. It would also be beneficial to create a glossary that compares different sustainability-related texts to identify commonly used words or phrases, which could lead to lexical standardisation. In addition to this, following the examples of other studies, it is recommended to involve people with cognitive disabilities in user tests to prototype content about climate change, ideally involving them from the outset of such initiatives.

Declarations

References

1. United Nations: Google teams up with UN for verified climate information. https://www.un.org/en/climatechange/google-search-information (2023). Accessed 29 March 2023
2. UNE: Norma española experimental UNE 153101 EX. Lectura fácil. Pautas y recomendaciones para la elaboración de documentos [Spanish experimental standard UNE 153101 EX. Easy Read. Guidelines and recommendation for the elaboration of documents] (2018)
3. Engelman, A., Craig, L., Iles, A.: Global disability justice in climate disasters: mobilizing people with disabilities as change agents. Health Aff. **41**, 1496–1504 (2022). https://doi.org/10.1377/hlthaff.2022.00474
4. Shriver-Rice, M., Vaughan, H.: What is environmental media studies? J. Environ. Media **1**, 3–13 (2020). https://doi.org/10.1386/jem_00001_2
5. Department of Agriculture, E., R.A.: Discussion Document on a Northern Ireland Climate Change Bill. Department of Agriculture, Environment and Rural Affairs, Belfast (2020)
6. Department of Agriculture, E., R.A.: Discussion Document on a Northern Ireland Climate Change Bill (Easy Read Version). Department of Agriculture, Environment and Rural Affairs, Belfast (2020)
7. Greater London Authority: London Environment Strategy: Executive Summary. Mayor of London, London (2018)
8. Greater London Authority: London Environment Strategy: Executive Summary (Easy Read Version). Mayor of London, London (2018)
9. Jodoin, S., Lofts, K., Bowie-Edwards, A., Leblanc, L., Rourke, C.: Disability Rights in National Climate Policies: Status Report (2022)
10. United Nations Environment Programme: The Paris Agreement (2015)
11. United Nations: Convention on the Rights of Persons with Disabilities. Department of Economic and Social Affairs Disability (2008)
12. Stein, P.J.S., Stein, M.A., Groce, N., Kett, M.: The role of the scientific community in strengthening disability-inclusive climate resilience. Nat. Clim. Chang. **13**, 108–109 (2023). https://doi.org/10.1038/s41558-022-01564-6
13. Sengel, Z.: How meaningful is 'can anybody hear me?' during a disaster? (2023)
14. Frieden, L.: The Impact of Hurricanes Katrina and Rita on People with Disabilities: A Look Back and Remaining Challenges (2006)
15. ISO: ISO/IEC 23859-1. Information technology—User interfaces—Part 1: Guidance on making written text easy to read and easy to understand. https://www.iso.org/obp/ui/#iso:std:iso-iec:23859:-1:dis:ed-1:v1:en (2023)
16. Federal Communications Commission: Cognitive Disabilities. https://www.fcc.gov/cognitive-disabilities
17. IFLA (International Federation of Library Associations and Institutions): Guidelines for easy-to-read materials. https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/publications/professional-report/120.pdf (2010)

18. Centre for Inclusive Design: Easy English versus Plain English. A guide to creating accessible content. Centre for Inclusive Design (2020)

19. Lindholm, C., Vanhatalo, U. (eds.): Handbook of Easy Languages in Europe. Frank & Timme, Berlin (2021)

20. González-Sordé, M., Matamala, A.: Empirical evaluation of Easy Language recommendations: a systematic literature review from journal research in Catalan, English, and Spanish. Univers. Access Inf. Soc. (2023). https://doi.org/10.1007/s10209-023-00975-2

21. Sutherland, R.J., Isherwood, T.: The evidence for easy-read for people with intellectual disabilities: a systematic literature review. J. Policy Pract. Intellect. Disabil. 13, 297–310 (2016). https://doi.org/10.1111/jppi.12201

22. Fuchs, J.: Leichte Sprache auf dem Prüfstand. Realisierungsvarianten von kausalen Relationen in Leichte-Sprache-Texten. Sprachwissenschaft 44, 441–480 (2019)

23. Deilen, S., Hansen-Schirra, S., Garrido, S.H., Maaß, C., Tardel, A. (eds.): Emerging Fields in Easy Language and Accessible Communication Research. Frank & Timme GmbH, Berlin (2023)

24. Hansen-Schirra, S., Maaß, C. (eds.): Easy Language Research: Text and User Perspectives. Frank & Timme, Berlin (2020)

25. Oliver, M.: Understanding Disability. From Theory to Practice. Macmillan Education UK, London (1996)

26. Chinn, D., Buell, S.: Easy Language in the UK. In: Lindholm, C., Vanhatalo, U. (eds.) Handbook of Easy Languages in Europe, pp. 623–657. Frank & Timme, Berlin (2021)

27. Dash, N.S., Arulmozi, S.: Limitations of Language Corpora. In: History, Features, and Typology of Language Corpora, pp. 259–272. Springer Singapore, Singapore (2018)

28. Arias-Badia, B., Matamala, A.: Audio description from an easy-to-understand language perspective: A corpus-based study in Catalan. Jostrans: J. Spec. Translat. (2023)

29. Department of Agriculture, E., R.A.: Discussion Document on a Climate Change Bill. https://www.daera-ni.gov.uk/consultations/climatechangediscussion

30. Baker, P., Hardie, A., McEnery, T.: A Glossary of Corpus Linguistics. Edinburgh University Press, Edinburgh (2006)

31. European Commission: How to write clearly. Publications Office (2012)

32. Accessible Information Working Group: Make it Easy: A guide to preparing Easy to Read Information. https://inclusionireland.ie/wp-content/uploads/2020/10/makeiteasyguide2011.pdf (2011)

33. Oakes, M.P.: Describing a translational corpus. In: Oakes, M.P., Ji, M. (eds.) Quantitative Methods in Corpus-Based Translation Studies: A practical guide to descriptive translation research, pp. 115–148. John Benjamins, Amsterdam (2012)

34. Perego, E.: Accessible Communication: A Cross-country Journey (2020)

35. Yaneva, V.: Easy-read documents as a gold standard for evaluation of text simplification output. In: Proceedings of the Student Research Workshop associated with RANLP 2015, pp. 30–36 (2015)

36. Knight, E., Blacher, J., Eisenhower, A.: Predicting reading comprehension in young children with autism spectrum disorder. School Psychology. 34, 168–177 (2019). https://doi.org/10.1037/spq0000277

37. Parker, C.E., Gorin, J., Bechard, S.: Adapting reading test items: Decreasing cognitive load to increase access for students with disabilities. In: Thurlow, M.L., Lazarus, S.S., Bechard, S. (eds.) Lessons learned in federally funded projects that can improve the instruction and assessment of low performing students with disabilities, pp. 17–58. University of Minnesota, National Center on Educational Outcomes, Minneapolis (2013)

38. Bond, S.: Gunning Fog Index. http://gunning-fog-index.com/index.html (2023). Accessed 5 March 2023

39. Readability Formulas: Automatic Readability Checker. https://readabilityformulas.com/free-readability-formula-tests.php (2023). Accessed 6 Feb 2023

40. Gunning, R.: The Technique of Clear Writing. McGraw-Hill, New York (1952)

41. SketchEngine: enTenTen: Corpus of the English Web. https://www.sketchengine.eu/ententen-english-corpus/ (2020)

42. García Muñoz, Ó.: Lectura fácil: métodos de redacción y evaluación. Real Patronato sobre Discapacidad, Madrid (2012)

43. Inclusion Europe: Information for all: European standards for making information easy to read and understand. https://inclusion-europe.eu/wp-content/uploads/2015/03/2113-Information-for-all-16.pdf (2009)

44. Rink, I.: Rechtskommunikation und Barrierefreiheit. Zur Übersetzung juristischer Informations- und Interaktionstexte in Leichte Sprache. Frank & Timme, Berlin (2020)

45. New, B., Ferrand, L., Pallier, C., Brysbaert, M.: Reexamining the word length effect in visual word recognition: new evidence from the English Lexicon Project. Psychon. Bull. Rev. 13, 45–52 (2006). https://doi.org/10.3758/bf03193811

46. Bochkarev, V.V., Shevlyakova, A.V., Solovyev, V.D.: The average word length dynamics as an indicator of cultural changes in society. Soc. Evol. Hist. 14, 153–175 (2015)

47. Perego, E.: Into the language of museum audio descriptions: a corpus-based study. Perspectives: Stud. Translat. Theory Pract. 27, 333–349 (2019). https://doi.org/10.1080/0907676X.2018.1544648

48. ILSMH-EA, (International League of Societies for the Mentally Handicapped): Make it Simple. European Guidelines for the Production of Easy-to-Read Information for People with Learning Disability for authors, editors, information providers, translators and other interested persons

49. Johansson, V.: Lexical diversity and lexical density in speech and writing: a developmental perspective. Working papers/Lund University, Department of Linguistics and Phonetics 53, 61–79 (2008)

50. The Writing Clinic: Clear Writing: How to Achieve and Measure Readability. http://notorc.blogspot.com/2006/09/devils-in-details-measuring.html

51. European Commission: Clear writing for Europe. https://commission.europa.eu/about-european-commission/departments-and-executive-agencies/translation/clear-writing-europe_en (2023). Accessed 30 March 2023