

The Simplification of the Language of Public Administration: The Case of Ombudsman Institutions

Gabriel González-Delgado, Borja Navarro-Colorado

University of Alicante

Alicante, Spain

ggd7@gcloud.ua.es, borja@dlsi.ua.es

Abstract

Language produced by Public Administrations has crucial implications in citizens' lives. However, its syntactic complexity and the use of legal jargon, among other factors, make it difficult to be understood for laypeople and certain target audiences. The NLP task of Automatic Text Simplification (ATS) can help to the necessary simplification of this technical language. For that purpose, specialized parallel datasets of complex-simple pairs need to be developed for the training of these ATS systems. In this position paper, an on-going project is presented, whose main objectives are (a) to extensively analyze the syntactical, lexical, and discursive features of the language of English-speaking ombudsmen, as samples of public administrative language, with special attention to those characteristics that pose a threat to comprehension, and (b) to develop the OmbudsCorpus, a parallel corpus of complex-simple supra-sentential fragments from ombudsmen's case reports that have been manually simplified by professionals and annotated with standardized simplification operations. This research endeavor aims to provide a deeper understanding of the simplification process and to enhance the training of ATS systems specialized in administrative texts.

Keywords: text simplification, public administrative language, parallel corpus

1. Introduction

Legal language, when it addresses laypeople, may be difficult to be understood. This lack of understanding in asymmetrical communication between experts and non-experts may lead to negative consequences in people's lives. Within the legal domain, administrative language is the one citizenship has the most relation with. It is the language produced by public bodies for the implementation of laws and legal regulations. However, obscure structures and complex terminology can pose a threat to the comprehension of its meaning, preventing people from being able to complete administrative procedures. For that reason, there exist some civil movements, such as the Plain English campaign, that advocate for the right to be addressed in a clear and understandable way by Public Administrations.

Ombudsman institutions, as whistler-blowers and guarantors of citizens' rights against maladministration, can also play an evangelization role by putting forward good-practice recommendations, denouncing abusive practices, and training public servants in charge of writing this type of texts. Nevertheless, this is a costly and time-consuming task. It is in this context that automatic text simplification (ATS) can be of use to make technical language clearer and more comprehensible.

This paper is framed within one of the author's on-going research project. Its first main objective is to analyze the linguistic features of the language of ombudsman offices as an example of public administrative language. A deeper understanding of this specialized language can contribute to

advance in its necessary simplification. To that end, corpus linguistics enables the processing of large amounts of texts to obtain quantitative results. The choice of compiling a corpus out of texts produced by English-speaking ombudsmen (named the OmbudsCorpus) is not arbitrary. All the ombudsmen's case reports and annual reports are publicly available on their websites, making it an abundant source of linguistic evidence in the domain of administrative language. Besides, they are present in almost every English-speaking country, which allows for variational studies. The second objective of this project is the creation of a parallel corpus of original fragments from ombudsmen's case reports and their manually simplified version. These simplified fragments include standardized annotations on every simplification operation performed, so the parallel corpus can be used as reference data for the training and evaluation of ATS systems specialized in public administrative language.

This paper will be structured as it follows: Section 2 will discuss the main issues regarding the simplification of legal language. In Section 3, the different methodological frameworks for the creation of reference data for ATS systems will be explained. An account of the OmbudsCorpus, including its sources and the methodology followed for its development, will be provided in Section 4. Finally, some conclusions and insights on the contributions this research project aims at will be put forward in Section 5.

2. Issues about the simplification of legal language

The first and main issue about simplification is the notion of simplicity itself. When can an utterance

be considered complex or simple? At what point does a text stop being complex in the process of simplification? Is a complex text equally difficult to everyone? To be able to answer these questions, one must firstly approach the issue of complexity/simplicity as a continuum. We cannot say that a text is complex or simple *per se*, but that some of its components may entail some complexity and others that may contribute to an easy understanding. That is, the difficulty or ease of a text is conditioned by several internal and external factors. The familiarity of its lexicon or the syntactic complexity are some instances of internal factor than can influence comprehensibility. But even the complexity of these internal factors may be differently perceived from reader to reader. Thus, what ultimately determines the comprehensibility of a text is its target audience. Different addressee profiles (children, people with low literacy levels or lay-people, the elderly, foreigners, people with intellectual or speech disabilities, etc.) and with different backgrounds (for instance, familiarity with technical jargon and discursive genres) may present different needs to understand the content of legal documents (Garimella et al., 2022).

Legal language is characterized by the overuse of formulaic and archaic language (e.g. Latinisms), passivity and impersonal structures, abbreviations, non-finite clauses headed by gerunds, among others (see Alcaraz et al., 2013; Bhatia, 1987; Charrow et al. (2015); Danet, 1980, 1983, 1985; Gustafsson, 1983; Maley, 1987; Mellinkoff, 1963). All these features together result in dense and complicated texts that could be written in a more user-friendly manner while preserving its intended meaning. That is what plain language recommendations aim at. Throughout all the English-speaking countries that have joined this movement, it is possible to find the following ten common recommendations (see Section 8 for references to Plain Language manuals):

1. Keep your sentences short (between 15-20 words).
2. Use simple, clear words.
3. Avoid complex, technical words and choose a simpler synonym.
4. Take care when using foreign expressions, namely from French and Latin origin.
5. Take care when using initials and acronyms.
6. Avoid chains of nouns, also known as *nouns strings* (“nouns strung together to act as adjectives”).
7. Construct sentences following the order: Subject + Verb + Objects.
8. Use active voice instead of passive constructions and impersonality.
9. Address the receiver directly.

10. Consider using illustrations, tables and lists to make complex material easier to understand.

As it can be seen, these recommendations try to tackle some of the main features of legal language that make a text complex. However, some of their propositions are too vague and generic, and they fail to take into account some crucial factors that influence comprehensibility.

Simplification, whether in general or in technical contexts, is sometimes seen as the replacement of a long word for a shorter one, or the shortening of long sentences, but that does not necessarily lead to better comprehension (McNamara et al., 2014). In fact, according to Brysbaert et al. (2011), the variable ‘word length’ only correlates to 1.2% of the reading processing time. On the contrary, what really influences the degree of complexity of a word is its *frequency* (van Heuven et al., 2014). Other psycholinguistic parameters that have an impact on the lexical decision time (Brysbaert et al., 2011) are the degree of *concreteness* (Brysbaert et al., 2014) (also referred to as *sensimotor content* (Lynott et al., 2020)), *age of acquisition* (Kuperman et al., 2012), the *semantic density* (Hoffman et al., 2013), and the *local coherence* (Hoffman et al., 2018).

In the same vein, syntactic complexity is not just a matter of length. It can be better explained by the analysis of the frequency of certain Universal Dependencies (De Marneffe et al., 2021), as explored by Deilen et al. (2023): *acl* (adnominal clause or clausal modifier of noun), *advcl* (adverbial clause modifier), *ccomp* (clausal component), *csubj* (clausal subject), *xcomp* (open clausal element) or *parataxis* (parataxis relation).

For that reason, it is necessary to implement these variables when determining the degree of complexity of a text, as it will be shown in Section 4.3, so that the simplification of legal language, either manually performed by professional or automatized by a NLP tool, can produce objectively clearer and simpler outputs that take into consideration the subjective needs of the target population.

3. Datasets for text simplification and evaluation issues

3.1 Reference data

The lack of complex-simple parallel corpora developed from legal texts is one of the main problems for the task of ATS in this domain (Garimella et al., 2022). These parallel datasets are the reference data that ATS systems are trained on. Besides, to evaluate the performance of an ATS system, outputs need to be compared to that reference data (Cardon et al., 2022, p. 1842). Thus, the approach taken to determine

what reference data an ATS system will be trained with crucially impacts the outputs produced.

Various methodologies for the creation of reference data have been reviewed by Grabar and Saggion (2022). While *expert judgment* or content extracted from *textbooks* may be established as reference data, these methods are heavily reliant on the theoretical comprehension of the producers regarding the requirements of the target audience. To address this constraint, *crowd-sourced* simplifications are used to gather extensive reference data based on the target population's judgement. However, as an online process, it is difficult to fully verify whether contributors fit in that aimed audience. An alternative method involves the application of *eye-tracking*, wherein the eye movements of readers are monitored as they engage with a reference text, enabling the quantification of attention allocation. Prolonged fixation on specific lexical units indicates higher complexity. As a drawback, this approach demands meticulous control and technical support.

Annotated reference data curated by professionals appears as another prevalent technique. Human annotation enhances the efficacy of Automatic Text Simplification (ATS) systems, particularly for rule-based systems, by elucidating the intricacies of lexical, syntactic, and even pragmatic simplification processes. However, this approach needs substantial efforts and is subject to the limitations of time and resources. Moreover, it retains a subjective element influenced by annotators' comprehension of simplification rules (Shardlow, 2014).

Newsela (Xu et al., 2015; 1,130 sentences and 5 simplified versions per sentence) and TurkCorpus (Xu et al., 2016; 2,350 sentences with 8 simplified references each) are the main reference data produced by human simplification and annotation used for ATS evaluation. They are in English and do not focus on any specific domain. The ASSET_{ann} corpus (Cardon et al., 2022) has recently been proposed as an attempt to standardize the annotation process in the simplification task.

In other languages, it is possible to find the *Dsim* corpus (Klerke and Søggaard, 2012), in Danish, with roughly 50,000 sentences pairs simplified from news telegrams by trained journalists; in Brazilian Portuguese, Specia et al. (2008) crafted a manual based on the simplification and annotation of ca. 2,000 sentences extracted from news articles; in Japanese, see Goto et al. (2015), who combined automatic alignment for training data (~10,000 pairs) and manual alignment for validation (~700) and testing (~2,000); in Italian, *Terence* (Brunato et al., 2014) was developed for the simplification of texts targeting children and it contains approximately 1,000 manually aligned

pairs. *SIMPITIKI* (Tonelli et al., 2016) is another corpus in Italian compiled from Wikipedia, which contains 345 sentence pairs and 575 annotations of simplification operations. Battisti et al. (2020) presented a parallel corpus in German which included annotation on text structure, typography, and images. In this same language, Spring et al. (2021) reported their work on a corpus in which simplifications were classified within A1, A2, and B1 levels of the Common European Framework of Reference for Languages. In French, *CLEAR* has progressively been developed (Cardon & Grabar, 2018, 2020; Koptient et al., 2019) as a specialized corpus in the biomedical domain with more than 4,500 parallel sentences in its latest version (2020). In Spanish, CLARA-MeD (Campillos-Llanos et al., 2022) is also a medical-domain corpus made up of about 25,000 pairs. EASIER (Alarcon et al., 2023) is a domain-independent corpus in Spanish with only lexical annotations.

For the task of ATS of legal documents, some specific corpora exist. *SimPA* (Scarton, et al., 2018) is a corpus in English extracted from the Sheffield City Council's website. Through crowdsourcing, it is made up of 1,100 original sentences with 3 lexically simplified versions and one syntactical simplified pair. *SIMPITIKI* (Tonelli et al., 2016) also contains a defined selection of 591 simplified sentences from the Public Administration domain that were manually created and annotated.

3.2 Evaluation

The optimal approach for evaluating the performance of Automated Text Simplification (ATS) systems is through human assessment, which can be conducted either by expert linguists or by a diverse sample from the target population (Alva-Manchego et al., 2020). Within this methodology, evaluators typically employ a Likert scale ranging from 1 to 5 to rate outputs across three key criteria, namely *fluency* (grammatical correctness), *adequacy* (preservation of meaning), and *simplicity* (Štajner et al., 2016).

However, this method requires substantial human and time resources. Consequently, automatic evaluation metrics have been developed, with BLEU (Papineni et al., 2002) and SARI (Xu et al., 2016) being the most widely used. It is essential to note, nonetheless, that these metrics have not escaped criticism (Grabar & Saggion, 2022), as their primary focus lies in measuring lexical similarity rather than simplicity.

Moreover, classical readability metrics, including the Flesch Reading Ease (Flesch, 1948), Gunning Fog Index (Gunning, 1952), Automatic Readability Index (ARI) (Senter and Smith, 1967), and particularly Flesch-Kincaid Grade Level (Flesch, 1975) are employed to automatically evaluate ATS systems (Alva-Manchego et al., 2020). Nevertheless, there has been significant

criticism regarding the use of these metrics to gauge text simplicity, as they predominantly consider factors such as word and sentence length (Crossley et al., 2008), which are deemed but superfluous factors to simplicity. Traditional readability metrics do not encompass psycholinguistic factors that contribute to text complexity. McNamara et al. (2014) observe a strong correlation between psycholinguistic features such as word frequency, familiarity, age of acquisition, concreteness, and imageability, and lexical decision time, suggesting they offer a more truthful measure of lexical complexity.

4. The OmbudsCorpus

4.1 Corpus compilation

The corpus that has been compiled for the characterization of this language of the public administration is composed of Annual Reports and Case Reports from English-speaking ombudsmen (Appendix A provides a detailed list of the country these institutions are from). Texts' date of production ranges from 1992 to 2022. This information is annotated in each document so variational factors in terms of diachronic and diatopic variation may also be explored.

Ombudsman offices assign complains to an area and then publish on their websites the result in a case report. To allow for a homogeneous composition, texts were selected from three thematic areas which are shared across all Offices: Education, Health, and Housing.

Besides, the volume of workload in each ombudsman is different mainly due to demographic reasons, and so is the amount of available documentation. If all samples from all the ombudsmen were analyzed at once in a single corpus, to guarantee the representativeness of all the sources, the proportion of words per ombudsman would be limited to the one with less available information. Therefore, different subcorpora including the maximum amount of information within each country, while keeping each area proportionally represented as far as size is concerned, have been established. Thus, the linguistic analysis will be performed separately for each subcorpus, and results will be compared among them to extract common features. The table in Appendix A also includes figures on the number of tokens per country and thematic area. The overall size of the OmbudsCorpus is ca. 12,600,000 tokens (~11.7M from Annual Reports and ~950K from Case Reports).

4.2 Corpus simplification and annotation

The parallel OmbudsCorpus is composed of original fragments from case reports and its simplified counterpart. The simplification was performed by two professionals (expert linguists in the field of simplification of languages for

specific purposes), who also included the annotation of each transformation operation that fragments had undergone to be rendered simpler.

To select the original fragments, each text was analyzed in terms of complexity. To determine lexical complexity, the variables of 'word frequency', 'familiarity', 'concreteness', and 'imageability' of content words were measured by means of TAALES 2.2 tool (Kyle et al., 2018). The variables under consideration in measuring syntactic complexity were 'subordinating conjunctions per clause' (mark_per_cl), 'passive auxiliary verbs per clause' (auxpass_per_cl), 'dependents per clause' (cl_av_deps), and 'clausal complements per clause' (ccomp_per_cl). TAASSC 1.3.8 (Kyle, 2016) was used for that purpose. Even though some scholars (see Alva-Manchego et al., 2020, p. 40; Crossley et al., 2008) advice against readability metrics to assess actual complexity, the Flesch-Kincaid grade level, as a well-established metric, was applied to allow for comparison with other state-of-the-art datasets which include it.

The most complex texts according to these metrics were selected, so the resulting simplification and annotation will present the most paradigmatic instances. These fragments contain more than one sentence, so simplification was performed at a supra-sentential level. Thus, the limitation of evaluation measures only being based at the level of sentence (Todirascu et al., 2013) is meant to be overcome. Almost half of these texts had been produced by the Scottish Public Service Ombudsman. Regarding thematic areas, Housing is the most complex one. In terms of the date of publication, the vast majority of them belong to the last decade.

Texts simplified include the type of simplification applied. Only simplified fragments were annotated to reflect the different simplification operations they had undergone in comparison to the original fragment. It has been represented with XML tags (see Appendix B) following the formalization proposed by Cardon et al. (2022). The main tags correspond to common general operations: insert, delete, replace. Each one has its own subtypes, as insert or delete modifiers or replace with synonyms. For other actions there are also specific tags, as <move> or <verbf/> (when there is a modification of a verbal feature, like tense or modality). Finally, some specific tags have been defined for very common specific actions such as "to personal form" (<fromImp/>). In so doing, the recommendations of Plain Language can be translated to standardized simplification operations, as presented in Appendix C.

4.3 Evaluation

The OmbudsCorpus is evaluated at different instances. Regarding the annotation of the

simplification operations, a parallel comparison of the tags used in each fragment by both annotators will be performed. It is important to bear in mind, as Stodden & Kallmeyer (2022) warn, that disagreement “does not always indicate a bad quality of the annotations, (...) it can be due to different subjective perspectives on the task”. In fact, because there is no “perfect” simplification, two humans can create different simplified texts from the same textual source, both correct. For this reason, we think that the evaluation of simplification should be framed within the perspectivist approach to corpus annotation (Cabitza et al., 2023). This approach considers that the disagreement between two annotators is not an error, but rather different visions (or interpretations) of the same phenomenon, both correct.

As far as the assessment of the simplified versions is concerned, the same complexity variables as the ones applied to the original fragments (see Section 4.2) are analyzed to establish the extent of the simplification. T-tests are performed for each parameter to compare if there is a statistically significant improvement (p-value <0.05).

Metric	Original	Simp	p-value
KF_Freq_CW_Log	2.142	2.301	< 0.05
TL_Freq_CW_Log	2.785	2.983	< 0.05
Brown_Freq_CW_Log	1.299	1.494	< 0.05
Familiarity	559.43	565.87	< 0.05
Concreteness	345.59	342.66	> 0.05
Imageability	368.14	370.48	> 0.05
ccomp_per_cl	0.216	0.164	< 0.05
mark_per_cl	0.210	0.179	> 0.05
auxpass_per_cl	0.126	0.084	< 0.05
Flesch-Kincaid Grade Level	14.960	8.857	< 0.05

Table 1: Comparison of metrics between original and simplified versions.

All metrics improved in the simplified version, except for *concreteness*. The replacement of complex words has been done by more frequent and familiar words, as the metrics on logarithmic frequencies and the *familiarity* metric indicate. All of them with a p-value <0.05. However, the lexicon chosen for the substitution of complex words still retains high levels of abstraction. *Imageability*, which usually correlates with concreteness, shows some improvement, even though the difference is not statistically significant either.

Regarding the syntactic metrics analyzed, the average of clausal components per clause (*ccomp_per_cl*) and passive verbs per clause (*auxpass_per_cl*) was reduced significantly. In other words, simplified fragments contain fewer subordinate clauses and more sentences in the active voice. Despite the reduction in subordinate

clauses, the difference in the number of subordinate conjunctions per clause (*mark_per_cl*) is not statistically significant.

This analysis allows us to identify specific pairs of fragments within the parallel corpus which may require further simplification so that an optimal simplification may be reached.

5. Conclusions

In this paper, we have presented the main objectives of this on-going research project and the research needs it targets. Previous to the task of the simplification of the public administrative language, it is necessary to know the stylistic features that are present in this register and that convey the most complexity to citizenship. The compilation and analysis of a specialized corpus from ombudsmen’s text will fill in this knowledge gap.

Regarding the notion of simplicity itself, it is necessary to approach this issue from the concept of comprehensibility, instead of that of readability, as it is often done. As it has been explained, quantitative indices such word or sentence length cannot determine by themselves the complexity of a text. Psycholinguistic studies on the parameters influencing comprehension and more sophisticated metrics on syntactic structures can shed some light on this regard. These are the metrics that have been implemented in the evaluation of the OmbudsCorpus. It is important to bear in mind that the psycholinguistic parameters included in the tool TAALES (i.e., familiarity, concreteness, age of acquisition, etc.) are based on human ratings. That is where the key to determine simplicity/complexity lies.

Literature on automatic text simplification of specialized domains highlights the need for the creation of parallel datasets that serve as reference data for the training of ATS systems. Annotated reference data have proved to achieve the best state-of-the-art results. The parallel OmbudsCorpus has been developed following this methodology, incorporating the annotation of all the simplification operations applied to the original fragment. It is composed of supra-sentential pairs, in an attempt to overcome the limitations of previous datasets which remain at sentential level. It has also been enriched with syntactic and lexical parameters so the degree of complexity can objectively be compared from the original fragments to its simplified version. The intended enrichment with ratings by target audiences is an additional measure that would definitely establish a benchmark in the validation and assessment of reference data in the legal domain. A test with different ATS systems will determine the usefulness of all this annotated information in a parallel corpus.

6. Acknowledgements

This paper has been partially funded by the Spanish Government through the R&D projects “CORTEX: Conscious Text Generation” (PID2021-123956OB-I00, funded by MCIN/AEI/10.13039/501100011033/ and by “ERDF A way of making Europe”) and “CLEAR.TEXT: Enhancing the modernization public sector organizations by deploying Natural Language Processing to make their digital content CLEARER to those with cognitive disabilities” (TED2021-130707B-I00), and by the Generalitat Valenciana through the project “NL4DISMIS: Natural Language Technologies for dealing with dis- and misinformation with grant reference (CIPROM/2021/21)”.

7. Bibliographical References

- Alarcon R, Moreno L, Martínez P (2023) EASIER corpus: A lexical simplification resource for people with cognitive impairments. *PLoS ONE*, 18(4).
<https://doi.org/10.1371/journal.pone.0283622>
- Alcaraz, E., Campos, M. A. y Miguélez, C. (2013). *El inglés jurídico norteamericano*. Barcelona, Ariel.
- Alva-Manchego, F., Scarton, C., & Specia, L. (2020). Data-driven sentence simplification: Survey and benchmark. *Computational Linguistics*, 46(1), 135-187.
- Battisti, A., Pfütze, D., Säuberli, A., Kostrzewa, M., & Ebling, S. (2020). A Corpus for Automatic Readability Assessment and Text Simplification of German. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 3302–3311), Marseille, France. European Language Resources Association.
- Bhatia, V. K. (1987). Language of the law. *Language Teaching*, 20(4), 227–234.
<https://doi.org/10.1017/S026144480000464X>
- Bott, S., & Saggion, H. (2014). Text simplification resources for Spanish. *Language Resources and Evaluation*, 48(1), 93-120.
- Brunato, D., Dell’Orletta, F., Venturi, G., & Montemagni, S. (2014). Defining an annotation scheme with a view to automatic text simplification. *Defining an annotation scheme with a view to automatic text simplification* (pp. 87-92).
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect. *Experimental psychology*.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, 46, 904-911.
- Cabitza, F., Campagner, A., Basile, V. (2023). Toward a Perspectivist Turn in Ground Truthing for Predictive Computing, *Proceedings of the AAAI Conference on Artificial Intelligence*.
- Campillos Llanos, L., Terroba Reinares, A. R., Zakhir Puig, S., Valverde, A., & Capllonch-Carrión, A. (2022). Building a comparable corpus and a benchmark for Spanish medical text simplification.
- Cardon, R., & Grabar, N. (2020). French biomedical text simplification: When small and precise helps. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 710-716).
- Cardon, R., Bibal, A., Wilkens, R., Alfter, D., Norré, M., Müller, A., Watrin, P., & François, T. (2022). Linguistic Corpus Annotation for Automatic Text Simplification Evaluation. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (pp. 1842-1866).
- Charrow, V. R., Crandall, J. A., & Charrow, R. P. (2015). Characteristics and functions of legal language. In *Sublanguage* (pp. 175-190). de Gruyter.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37-46.
- Coster, W., & Kauchak, D. (2011, June). Simple English Wikipedia: a new text simplification task. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 665-669).
- Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *Tesol Quarterly*, 42(3), 475-493.
- Danet, B. (1980). Language in the legal process. *Law & Soc’y Rev.*, 14, 445.
- Danet, B. (1983): Language in legal and bureaucratic settings. In: *Language as a social problem, special issue of Transaction-Society* (A. Grimshaw, ed.).
- Danet, B. (1985). Legal discourse. *Handbook of discourse analysis*, 1, 273-291.
- De Marneffe, M. C., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal dependencies. *Computational linguistics*, 47(2), 255-308.
- Deilen, S., Garrido, S. H., Lapshinova-Koltunski, E., & Maaß, C. Using ChatGPT as a CAT tool in Easy Language translation.
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221.
- Garimella, A., Sancheti, A., Aggarwal, V., Ganesh, A., Chhaya, N., & Kambhatla, N. (2022). Text Simplification for Legal Domain: Insights and Challenges. In *Proceedings of the Natural Legal Language Processing Workshop 2022* (pp. 296-304).
- Goto, I., Tanaka, H., & Kumano, T. (2015). Japanese news simplification: tak design, data set construction, and analysis of simplified text.

- In *Proceedings of Machine Translation Summit XV: Papers*.
- Grabar, N. & Cardon, R. (2018). CLEAR – simple corpus for medical French. In *Proceedings of the 1st Workshop on Automatic Text Adaptation (ATA)* (pp. 3–9), Tilburg, the Netherlands. Association for Computational Linguistics.
- Grabar, N., & Saggion, H. (2022). Evaluation of Automatic Text Simplification: Where are we now, where should we go from here. In *Traitement Automatique des Langues Naturelles* (pp. 453-463). ATALA.
- Gunning, R. (1952). "The technique of clear writing". Information Transfer and Management. McGraw-Hill.
- Gustafsson, M. (1984). The syntactic features of binomial expressions in legal English. *Text-Interdisciplinary Journal for the Study of Discourse*, 4(1-3), 123-142.
- Hoffman, P., Jefferies, E., Haffey, A., Littlejohns, T., & Lambon Ralph, M. A. (2013). Domain-specific control of semantic cognition: A dissociation within patients with semantic working memory deficits. *Aphasiology*, 27(6), 740-764.
- Hoffman, P., Cogdell-Brooke, L., & Thompson, H. E. (2020). Going off the rails: Impaired coherence in the speech of patients with semantic control deficits. *Neuropsychologia*, 146, 107516.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). Flesch-kincaid grade level. *Memphis: United States Navy*.
- Klerke, S., & Søgaard, A. (2012, May). DSIm, a Danish Parallel Corpus for Text Simplification. In *Proceedings of the Eight International Conference on Language Resources and Evaluation, LREC'12* (pp. 4015-4018).
- Koptient, A., Cardon, R., & Grabar, N. (2019). Simplification-induced transformations: typology and some characteristics. In *BioNLP 2019*.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior research methods*, 44, 978-990.
- Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior research methods*, 52, 1271-1291.
- Maley, Y. (1987). *The Language of Legislation*. 16(1), 25–48.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- Mellinkoff, D. (1963). *The Language of the Law*. Little, Brown.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics* (pp. 311-318).
- Scarton, C., Paetzold, G., & Specia, L. (2018, May). Simpa: A sentence-level simplification corpus for the public administration domain. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (pp. 4333-4338).
- Senter, R. J., & Smith, E. A. (1967). *Automated readability index* (pp. 1-14). Technical report, DTIC document.
- Shardlow, M. (2014). A Survey of Automated Text Simplification. *International Journal of Advanced Computer Science and Applications*, 4(1), 58-70.
- Siddharthan, A. (2014). A survey of research on text simplification. *ITL-International Journal of Applied Linguistics*, 165(2), 259-298.
- Specia, L., Aluísio, S. M., & Pardo, T. A. (2008). Manual de simplificação sintática para o português. *NILC, Sao Carlos-SP*.
- Spring, N., Rios, A., & Ebling, S. (2021). Exploring German multi-level text simplification. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP)* (pp. 1339–1349).
- Štajner, S., Popovic, M., Saggion, H., Specia, L., & Fishel, M. (2016). Shared task on quality assessment for text simplification. In *Proceedings of the Workshop on Quality Assessment for Text Simplification - LREC 2016* (pp. 22–31).
- Stodden, R., & Kallmeyer, L. (2022, May). TS-ANNO: an annotation tool to build, annotate and evaluate text simplification corpora. In *Proceedings of the 60th annual meeting of the association for computational linguistics: system demonstrations* (pp. 145-155).
- Todirascu, A., François, T., Gala, N., Fairon, C., Ligozat, A. L., & Bernhard, D. (2013, October). Coherence and cohesion for the assessment of text readability. In *Proceedings of 10th International Workshop on Natural Language Processing and Cognitive Science (NLPCS 2013)* (pp. 11-19).
- Tonelli, S., Aproso, A. P., & Saltori, F. (2016). SIMPITIKI: a Simplification corpus for Italian. In *CLiC-it/EVALITA* (pp. 4333-4338).
- Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology*, 67(6), 1176-1190.
- Xu, W., Callison-Burch, C., & Napoles, C. (2015). Problems in current text simplification research: New data can help. *Transactions of the Association for Computational Linguistics*, 3, 283-297.
- Xu, W., Napoles, C., Pavlick, E., Chen, Q., & Callison-Burch, C. (2016). Optimizing statistical machine translation for text

simplification. *Transactions of the Association for Computational Linguistics*, 4, 401-415.

Zhu, Z., Bernhard, D., & Gurevych, I. (2010). A monolingual tree-based translation model for sentence simplification. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)* (pp. 1353-1361).

8. Language Resource References

Plain English manuals

Australia Government Office of Parliamentary Counsel. (1993). *Plain English Manual*, 1–40.

European Commission. (n.d.). *Claire's Clear Writing Tips*.

National Adult Literacy Agency. (n.d.). *Plain English guidelines at a glance*, 1–2.

NWT Literacy Council. (2013). *A Plain Language Handbook: Write for your Reader*. <https://doi.org/10.1109/tpc.1980.6501907>

Plain English Campaign. (2001). *The A to Z of Alternative Words*.

Plain English Campaign. (2018). *How to write in plain English*. Retrieved from <https://www.plainenglish.co.uk/files/howto.pdf>

Plain Language Action and Information Network. (2011). *Federal Plain Language Guidelines*. Retrieved from <https://www.plainlanguage.gov/howto/guidelines/FederalPLGuidelines/FederalPLGuidelines.pdf>

Software

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* (Doctoral Dissertation). Retrieved from https://scholarworks.gsu.edu/alesl_diss/35/

Kyle, K., Crossley, S. A., & Berger, C. (2018). The tool for the analysis of lexical sophistication (TAALES): Version 2.0. *Behavior Research Methods* 50(3), pp. 1030-1046. <https://doi.org/10.3758/s13428-017-0924-4>

Stodden, R., & Kallmeyer, L. (2022, May). TS-ANNO: an annotation tool to build, annotate and evaluate text simplification corpora. In *Proceedings of the 60th annual meeting of the association for computational linguistics: system demonstrations* (pp. 145-155).

9. Appendices

Appendix A: Figures of the OmbudsCorpus by countries and types of texts (annual and case reports) and thematic area (Education, Health, and Housing)

Country*	Type	Ed	He	Ho	T
Australia	Annual	-	-	-	6.7M
	Case	155K	64K	124K	343K
Canada	Annual	-	-	-	1.9M
	Case	3.5K	3.3K	3.5K	10.5K
Ireland	Annual	-	-	-	588K
	Case	6.5K	6.9K	6.3K	19.8K
New Zealand	Annual	-	-	-	743K
	Case	17K	13.5K	15K	45.5K
UK	Annual	-	-	-	1.2M
	Case	198K	179K	139K	516K
USA	Annual	-	-	-	526K
	Case	4.2K	3.9K	4K	12K
TOTAL	Annual	-	-	-	11.7M
	Case	384K	272K	292K	948K
TOTAL					12.6M

* Texts for each country have been retrieved from the following sources:

- Australia:
 - [New South Wales Ombudsman](#).
 - [Northern Territory Ombudsman](#).
 - [Queensland Ombudsman](#).
 - [Tasmania Ombudsman](#).
 - [Victorian Ombudsman](#).
 - [Western Australia Ombudsman](#).
- Canada:
 - [British Columbia Ombudsperson](#).
 - [Manitoba Ombudsman](#).
 - [Ombud New Brunswick](#).
 - [Saskatchewan Ombudsman](#).
- Ireland:
 - [Ombudsman of Ireland](#).
- New Zealand:
 - [Ombudsman New Zealand](#).
- UK:
 - [Local Government and Social Care Ombudsman](#) (England).
 - [Northern Ireland Public Services Ombudsman](#).
 - [Parliamentary and Health Service Ombudsman](#) (UK).
 - [Public Services Ombudsman for Wales](#).
 - [Scottish Public Services Ombudsman](#).
- USA:
 - [Hawaii State Ombudsman](#).
 - [Iowa Office of Ombudsman](#).

Appendix B: List of tags used for annotation

Action	Tag
Delete proposition	<delete type="prop"/>
Delete modifier	<delete type="mod"/>
Delete for consistency	<delete type="cst"/>
Delete other	<delete type="other" subtype="..."/>
Replace with synonym (word-to-word)	<replace type="synonym" subtype="w2w">...</replace>
Replace with synonym (word-to-phrase)	<replace type="synonym" subtype="w2ph">...</replace>
Replace with synonym (phrase-to-word)	<replace type="synonym" subtype="ph2w">...</replace>
Replace with synonym (phrase-to-phrase)	<replace type="synonym" subtype="ph2ph">...</replace>
Replace with hypernym	<replace type="hypernym">...</replace>
Replace with hyponym	<replace type="hyponym">...</replace>
Replace segment with a pronoun	<replace type="pron">...</replace>
Replace singular with plural	<replace type="s2p">...</replace>
Replace plural with singular	<replace type="p2s">...</replace>
Modify verbal features	<verbf/>
Active to passive	<replace type="a2p">...</replace>
Passive to active	<replace type="p2a">...</replace>
Part-of-speech change	<POSchange/>
Split	<split/>
Merge	<merge/>
To impersonal form	<toImp/>
To personal form	<fromImp/>
Affirmation to negation	<replace type="a2n"/>...</replace>
Negation to affirmation	<replace type="n2a"/>...</replace>

Appendix C: "translation" of Plain Language recommendations to formalized simplification operations

Recommendation	Simplification operation
Eliminate unnecessary words or phrases	Delete modifier
	Delete proposition
Avoid complex words	Delete modifier
	Delete proposition
	Replace with synonym
	Replace with hypernym
	Replace with hyponym
Take care when using foreign expressions	Replace with synonym
	Specification
Use terms consistently throughout the text	Replace with synonym
	Insert for consistency
Avoid nominalization	Replace noun with verb
Keep sentences short	Delete modifier
	Delete proposition
	Split
	Merge
	Replace with synonym (phrase-to-word)
Use active voice instead of passivity and impersonality	Passive to active
	Modify verbal features
	To personal form
Use simple sentences: Subject + Verb + Complements	Delete modifier
	Delete proposition
	Delete for consistency
	Insert for consistency
Try to use affirmative sentences	Move
	Negation to affirmation
Address the receiver directly	Proximization