

# AMesure: a web platform to assist the clear writing of administrative texts

**Thomas François**  
IL&C, CENTAL/TeAMM  
UCLouvain

**Adeline Müller**  
IL&C, CENTAL  
UCLouvain

**Eva Rolin**  
IL&C, CENTAL  
UCLouvain

**Magali Norré**  
IL&C, CENTAL  
UCLouvain

## Abstract

This article presents the AMesure platform, which aims to assist writers of French administrative texts in simplifying their writing. This platform includes a readability formula specialized for administrative texts and it also uses various natural language processing (NLP) tools to analyze texts and highlight a number of linguistic phenomena considered difficult to read. Finally, based on the difficulties identified, it offers pieces of advice coming from official plain language guides to users. This paper describes the different components of the system and reports an evaluation of these components.

## 1 Introduction

In our current society, written documents play a central role as an information channel, especially in the context of communication between institutions and their target audiences (Madinier, 2009). Unfortunately, although efforts to raise the education level of the population worldwide have increased in recent decades, reports (OECD, 2016) point out that a significant proportion of citizens still have general reading difficulties. As regards administrative texts, various reading issues have been reported. For instance, Kimble (1992) reported that, in a survey carried out in the US, 58% of the respondents admitted to dropping out of an administrative process due to the reading difficulty.

Administrations have been aware of this issue for decades and have launched various initiatives to address it, the most prominent of which is the Plain Language movement. Plain language aims to increase the accessibility of legal documents for a general audience and has been shown to both reduce costs and please readers (Kimble, 1996). It has not only been promoted through various campaigns (e.g. *Plain English Campaign* in the UK) and writing guides (Gouvernement du Québec,

2006; Ministère de la Communauté française de Belgique, 2010; European Union, 2011; Plain Language Action and Information Network, 2011; Cutts, 2020), but also incorporated in some legal principles. However, its widespread application is still undermined due to, for example, the efforts required to train writers (Desbiens, 2008), or the necessity to persuade writers – especially legal ones – to abandon their flowery style, which is seen as a determinant of the image of expertise they project in the reader’s mind (Adler, 2012). This second reason, however, falls beyond the scope of the current study, which aims to address the first reason, i.e. writers’ training.

Recent research by Nord (2018) revealed that although several plain language guides are available to assist writers of administrative texts in their work, the guidelines provided in these guides are not always followed by writers, mainly because they are too vague and too numerous. To relieve writers from the need to keep all these guidelines in mind, we have designed a web platform, AMesure<sup>1</sup>, aimed at automatically identifying clear writing issues in administrative texts and providing simple writing advice that is contextually relevant. In its current state, the platform offers the three following functionalities: (1) providing an overall readability score based on a formula specialized to administrative texts; (2) identifying, in a text, linguistic phenomena that are assumed to have a negative effect on the comprehension of the text; (3) for the phenomena detected in step 2, proposing simplification advice found in plain language guides.

In the following sections, we first refer to some related work (Section 2), before describing the NLP analyses carried out to operate the system (Section 3.1). Then, we introduce the system and the way suggestions are provided (Section 3.2). The paper

<sup>1</sup>The platform is freely available online at <https://cental.uclouvain.be/amesure/>.

concludes with a report about the system performance (Section 4).

## 2 Related work

This work stands at the intersection between two very different fields: writing studies – “the interdisciplinary science that studies all the processes and knowledge involved in the production of professional writings and their appropriateness for the addressees” (Labasse, 2001) – and automatic text simplification (ATS), a branch of NLP that aims to automatically adapt difficult linguistic structures while preserving the meaning to enhance text accessibility.

Relevant facts from writing studies have already been covered in the introduction. As regards ATS, the last few years have witnessed the publication of numerous interesting studies, reviewed by Shardlow (2014), Siddharthan (2014), and Saggion (2017). In brief, the field has mainly focused on developing algorithms to automatically simplify complex words (*lexical simplification*) and/or complex syntactic structures (*syntactic simplification*). It has first relied on rule-based approaches (Chandrasekar et al., 1996; Siddharthan, 2011) in which a text is automatically parsed before being applied simplification rules defined by experts. Later, ATS has been assimilated to a translation task (the original version is translated into a simplified version) and addressed with statistical translation systems (Specia, 2010; Zhu et al., 2010). As neural machine translation has emerged under the impulse of deep learning, the Seq2Seq model has become prevalent for ATS too (Nisioi et al., 2017; Zhang and Lapata, 2017).

Some work has specifically focused on the issue of lexical simplification, which involves different techniques. Lexical simplification is generally operated in four steps, the first one being the identification of complex words. Some systems choose to consider all words as candidates for substitution (Bott et al., 2012); others use a list of complex words or machine learning techniques for classification of complex words (Alarcón et al., 2019). Once complex words have been identified, the next step is the generation of simpler synonyms for substitution, either by relying on lexical resources (De Belder and Moens, 2010; Billami et al., 2018), getting candidates from corpora (Coster and Kauchak, 2011), producing them with embeddings (Glavaš and Štajner, 2015; Paetzold

and Specia, 2016) or, more recently, with BERT (Qiang et al., 2020). In a next step, the candidates are semantically filtered to fit the context and are ranked according to their difficulty by classifiers using various word characteristics (e.g. frequency, embedding, morphemes, syllabic structures, etc.) (Paetzold and Specia, 2017; Billami et al., 2018; Qiang et al., 2020).

Although numerous ATS systems are described in publications, we have found only four of them that made their way through a web platform tailored to writers’ needs. Scarton et al. (2010) developed a simplification web platform for Portuguese, in which the user is able to either accept or reject simplifications done by the system. Similarly, Lee et al. (2016)’s system performs lexical and syntactic simplifications for English and supports human post-editing. More recently, Falkenjack et al. (2017) introduced TeCST, which is able to perform simplification at different levels, depending on the user. Finally, Yimam and Biemann (2018) implemented a semantic writing aid tool able to suggest context-aware lexical paraphrases to writers. None of these tools, however, have focused on writers of administrative texts, nor on French.

AMesure could also be related to the family of writing assistants, such as Word or LibreOffice. However, only a few of them provides writing advice based on specific criteria or plain language guides. There are some examples of these tools available for the general public in French: (1) Plainly<sup>2</sup>; (2) Lirec<sup>3</sup> which relies on the FALC guidelines, an equivalent of the Easy-to-Read language in French, tailored to people with a cognitive disability; or (3) Antidote<sup>4</sup>, which offers various writing advice to be clearer and includes five readability indexes. These are however commercial tools, whose scientific foundations are difficult to know and to compare to.

## 3 The platform

AMesure aims to help writers to produce clear and simple administrative texts for a general audience<sup>5</sup>.

For this purpose, it offers various diagnoses about the reading difficulty of a text as well as

<sup>2</sup><https://www.labrador-company.fr/outil-langage-clair/>

<sup>3</sup><http://sioux.univ-paris8.fr/lirec/>

<sup>4</sup><https://www.antidote.info/fr>

<sup>5</sup>People with low reading levels require even more simplified texts (with shorter sentences, no subordinated clauses, etc.). This “oversimplification” falls under the scope of the Easy Language domain.

advice on simpler ways of writing. Before moving to the description of the platform in Section 3.2, we first introduce the various NLP processes used to analyse the text and annotate difficulties in Section 3.1.

### 3.1 The analysis of the text

As soon as a text is uploaded on the platform, it is processed through various NLP tools to get a rich representation of the text, on which further rule-based processes are then applied. In a first step, the text is split into sentences and POS-tagged with MELT (Denis and Sagot, 2012), before being syntactically parsed with the Berkeley parser adapted for French (Candito et al., 2010). As a result, each sentence is represented as a dependency tree, on which we apply a set of handcrafted rules expressed in the form of regular expressions using the Tregex (Levy and Andrew, 2006) syntax. The rules currently implemented (François et al., 2018) are able to identify four classes of complex syntactic structures: passive clauses, relative clauses, object clauses, and adverbial clauses. Identifying these four classes is motivated by the characteristics of administrative texts. Passive clauses and infinitive verbs are often used in administrative texts to conceal the presence of the writer (Cusin-Berche, 2003), while other types of clause are used to provide the reader with as many detailed information as possible (Catherine, 1968). Parentheticals are also identified, as they are prone to hinder the reading process.

In a second step, the tagged text is further processed to carry out lexical analyses of the text. During this step, three types of lexical difficulties are identified. Firstly, rare words are detected relying on frequencies from Lexique3 (New et al., 2007), based on a threshold set empirically.

Secondly, technical terms are detected with some heuristics able to detect both simple terms and multi-word terms – a task that remains a challenge for current fully automatic approaches (da Silva Conrado et al., 2014) – that are included in a database. The database has been compiled from three different sources: (1) the official lists from the Belgian administration; (2) a list of terms extracted from a corpus of 115 administrative texts following the automatic extraction approach of Chung (2003) and then manually validated; and (3) a book describing various characteristics of the administrative style and listing administrative terms (Catherine, 1968). At the end of the collection phase, we

obtained 3,382 terms, some of which could, however, not be considered as difficult (e.g. academy, degree, jury, trainee, etc.). We therefore filtered the resource by excluding words found in the list of the 8000 simplest words in French (Gougenheim et al., 1964). As result, the final term database amounts to 2,481 entries.

Thirdly, abbreviations are automatically detected as they are known to produce reading errors, especially when they are used by specialized writers to communicate to non-specialized readers. For instance, Sinha et al. (2011) report that the Joint Commission on Accreditation of Healthcare Organizations estimated that 5% of medical errors are due to abbreviations. In our system, abbreviations are detected based on an abbreviation database, collected from Belgian public authorities. The database relate the extended version(s) of abbreviations (e.g. *communauté française, Institutions publiques de protection de la jeunesse*) with the corresponding abbreviated forms (e.g. *comm. fr.; IPPJ* and *I.P.P.J.* respectively). The list provided by public authorities was supplemented via a semi-automatic extraction process applied to our corpus of 115 administrative texts. This extraction process was based on manual rules maximizing the recall, in order to extract all forms prone to be abbreviations. Then, we filtered out all forms already in our list and manually checked the remaining ones, obtaining a final database with 2,022 entries.

### 3.2 Description of the platform

Leveraging the NLP analysis described above, the AMesure platform provides four types of diagnoses about texts to its users, as illustrated in Figure 1. The first diagnosis (marked by the letter A in the Figure 1) is a global readability score for the text. It is computed by a readability formula, specialized for administrative texts, that we previously developed (François et al., 2014). The output score ranges from 1 (for very easy texts) to 5 (for very complex texts) and is yielded by a support vector machine classifier combining 10 linguistic features of the text (e.g. word frequency, proportion of complex words, type-token ratio, mean length of sentence, ratio of past participle forms, etc.).

The second type of diagnosis (letter B in Figure 1) is more detailed and includes 11 readability yardsticks, each corresponding to one linguistic characteristic of the text known to affect reading. The psycholinguistic rationales for the choice of

these yardsticks have been discussed in length in François (2011), who has defined a set of 344 variables. Among this set, we have retained 11 yardsticks based on a correlational analysis (François et al., 2014). In the interface, the yardsticks are organised according to three linguistic dimensions of texts: lexicon, syntax, and textual aspects. The five lexical yardsticks capture (1) the percentage of difficult words, based on the list of 8000 simplest words in French (Gougenheim et al., 1964); (2) the number of rare words (see Section 3.1); (3) the density of abbreviations (see Section 3.1); (4) the proportion of unexplained abbreviations; and (5) the number of technical words (see Section 3.1). The four syntactic yardsticks include (1) the difficulty of the syntactic structures estimated roughly as the ratio of conjunctions and pronouns; (2) the mean number of words per sentence; (3) the ratio of structures considered as complex by plain language guides among all syntactic structures detected (see Section 3.1); and (4) the total number of sentences. As regards the two textual yardsticks, they include (1) a score corresponding to the level of personalization of the texts (text using pronouns at the first or at the second person are considered to be more readable (Daoust et al., 1996)); and (2) a score corresponding to the average intersentential coherence of the text. It is measured as the average cosine score between all adjacent sentences of the text, each of them being represented as a vector in a latent space (Foltz et al., 1998).

To render all these yardsticks more visual and more understandable, we project each of them on a five-degree scale, represented by colored feathers. The more feathers a yardstick gets, the more complex this linguistic dimension is supposed to be for reading. To transform the yardstick values into a five-degree scale, we applied the following method. Our corpus of 115 administrative texts has been annotated by experts on a five-degree difficulty scale (François et al., 2014). For each of our 11 yardsticks, we then estimated its Gaussian distribution (mean and standard deviation) on the corpus for each of the five levels. At running time, we simply compute the probability of the yardstick score for a given text to be generated by each of these five Gaussians and assign it the level corresponding to the higher probability.

The third type of diagnosis allows to directly visualize the text in which all complex phenomena annotated during the analysis step (see Section 3.1)

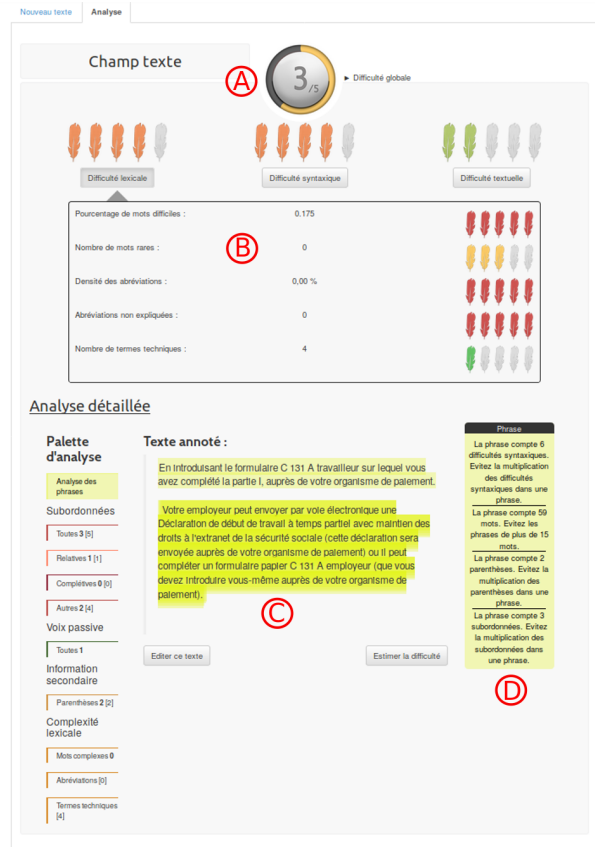


Figure 1: Result of a text analysis in AMesure.

are underlined, namely the three types of subordinated clauses, passives, parentheticals, rare words, abbreviations, and technical terms. For each of these categories, AMesure allows the user to select a tab showing only the respective phenomenon. It also offers a global view of the text in which complex sentences are highlighted in various shades of yellow (see letter C in Figure 1): the darker the yellow, the more difficult the sentence is to read.

Finally, the last functionality offers writing advice related to the complex phenomena detected (letter D in Figure 1). Two forms of advice are provided. On the one hand, we apply a list of 7 rules to filter out syntactic structures detected during the NLP analysis that should not be considered as complex. For instance, infinitive, participial, or even object clauses can be very short (e.g. *quand on décide d'avoir un bébé* or *le logement qu'il occupe*) and are therefore not at all a burden for reading. The filtering rules were defined based on writing guidelines from three plain language guides for French (Gouvernement du Québec, 2006; Ministère de la Communauté française de Belgique, 2010; European Union, 2011). We also extracted



from these guides some pieces of advice that are shown to users of the platform when a difficult syntactic phenomenon is detected. Examples of advice are: “This sentence has 50 words. Please avoid sentences longer than 15 words” or “This sentence has three subordinate clauses. Please avoid having so many subordinate clauses in a sentence”. On the other hand, we also offer simpler synonyms for words detected as rare words or technical words. The synonyms are taken from ReSyf (Billami et al., 2018), a lexical resource in which synonyms are ranked by difficulty. For now, we show the three simpler synonyms found in ReSyf for a given difficult word, letting the user to pick the best one. More advanced methods based on embeddings are, however, considered at the moment to improve the automatic selection.

#### 4 Evaluation of the system

To assess the performance of the various extraction algorithms included in our platform, three linguists manually annotated, in 24 administrative texts, the following five phenomena: passive structures, relative clauses, object clauses, adverbial clauses, and abbreviations<sup>6</sup>. The work of annotators was supported by guidelines focusing on difficult cases<sup>7</sup>. At the end of the annotation process, the expert agreement was evaluated using Fleiss’ kappa (see Table 1). The agreement was high for the rather easy tasks of annotating abbreviations and passive clauses. Detecting subordinate clauses is, however, a much more complex task, if only because it is also necessary to identify the type of structures. A common reference version was then built through consensus-building.

This gold-standard version of the annotation was manually compared to the output of AMesure for the 24 texts in the test set. Table 1 reports the results of this evaluation in terms of recall, precision, and F1-score for the different types of structures. Performance for the detection of passive clauses, relative clauses, adverbial clauses and abbreviations are satisfactory (F1 is above .8). By comparison, Zilio et al. (2017), who detect syntactic structures in English, obtained a precision of 0.88

<sup>6</sup>The detection of rare words and complex technical terms could not be assessed according to the same protocol as what matters is the psychological relevance of their identification for a given audience of readers. Further experiments with subjects are required to assess these two dimensions.

<sup>7</sup>For instance, infinitive clauses led by a semi-modal auxiliary such as *devoir* (ought to) or *pouvoir* (can) were discussed, as contradictory points of view can be found in grammars.

Phenomena	R	P	F1	$\kappa$
Passive clauses	0.92	0.92	0.92	0.92
Subordinated clauses (all)	0.84	0.87	0.85	0.47
Relative clauses	0.83	0.88	0.86	/
Object clauses	0.56	0.42	0.48	/
Adverbial clauses	0.78	0.83	0.8	/
Abbreviations	0.73	0.9	0.8	0.97
Total (macro-average)	0.83	0.9	0.86	0.79

Table 1: Recall (R), precision (P), F1, percentage of agreement and Fleiss’  $\kappa$  scores for the five phenomena detected in the platform.

and a recall of 0.62 for the relative clauses and a recall of 0.66 and a precision of 0.94 for infinitive clauses introduced by the particle “TO”. Chinkina and Meurers (2016) reached a recall of 0.83 and a precision of 0.71 for relative clauses. However, our system has trouble detecting object clauses, which have a F1-score of only 0.48. Investigation of the confusion matrix reveals that 77% of object clauses (37 out of 48) are correctly detected by AMesure, but 17 out of 37 are wrongly classified as adverbial clauses. This is a limited issue, as advice can still be provided even if the system gets the type of clause wrong.

#### 5 Conclusion

We have presented the AMesure system, which automatically analyzes the readability of French administrative texts based on classic readability metrics, but also on guidelines from plain language books. The system is freely available through a web platform and is aimed to help writers of administrative texts to produce more accessible documents and forms. To that purpose, it offers a global readability score for the texts, 11 readability yardsticks, a detailed diagnosis in which difficult linguistic words and syntactic structures are highlighted, and some plain language advice. We also carried out a manual evaluation of the system based on 24 administrative texts annotated by linguists. Performance is satisfactory, except as regards the identification of object clauses. More work is needed on this category, especially to distinguish it from adverbial clauses. We also plan to improve the system providing simpler synonyms by adding a semantic filter based on embedding models. Finally, we plan to conduct a study with real writers of administrative texts to measure the perceived usefulness of AMesure as a whole, but also the usefulness of each functionality.

## Acknowledgments

We would like to thank the "Direction de la langue française" from the Federation Wallonia-Brussels for its continued support to the AMesure project since 2014. We also want to acknowledge the wonderful work of Romain Pattyn, Gaëtan Ansolte, Baptiste Degryse on the interface and the great visuals of Brian Delmée.

## References

- M. Adler. 2012. The plain language movement. In *The Oxford handbook of language and law*.
- R. Alarcon, L. Moreno, I. Segura-Bedmar, and P. Martinez. 2019. Lexical simplification approach using easy-to-read resources. *Procesamiento de Lenguaje Natural*, 63:95–102.
- M. Billami, T. François, and N. Gala. 2018. ReSyf: A French lexicon with ranked synonyms. In *Proceedings of COLING 2018*.
- S. Bott, L. Rello, B. Drndarevic, and H. Saggion. 2012. Can Spanish Be Simpler? LexSiS: Lexical Simplification for Spanish. pages 357–374.
- M. Candito, J. Nivre, P. Denis, and E. H. Anguiano. 2010. Benchmarking of statistical dependency parsers for French. In *Proceedings of COLING 2010*, pages 108–116.
- R. Catherine. 1968. *Le style administratif*. A. Michel.
- R. Chandrasekar, C. Doran, and B. Srinivas. 1996. Motivations and methods for text simplification. In *Proceedings of the 16th conference on Computational Linguistics*, volume 2, pages 1041–1044.
- M. Chinkina and D. Meurers. 2016. Linguistically aware information retrieval: Providing input enrichment for second language learners. In *Proceedings of BEA 2016*, pages 188–198.
- T.M. Chung. 2003. A corpus comparison approach for terminology extraction. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 9(2):221–246.
- W. Coster and D. Kauchak. 2011. Simple English Wikipedia: A new text simplification task. In *Proceedings of ACL 2011*, pages 665–669. Association for Computational Linguistics.
- F. Cusin-Berche. 2003. *Les mots et leurs contextes*. Presses Sorbonne nouvelle.
- M. Cutts. 2020. *Oxford guide to plain English*. Oxford University Press, USA.
- F. Daoust, L. Laroche, and L. Ouellet. 1996. SATOCALIBRAGE: Présentation d'un outil d'assistance au choix et à la rédaction de textes pour l'enseignement. *Revue québécoise de linguistique*, 25(1):205–234.
- J. De Belder and M.-F. Moens. 2010. Text simplification for children. In *Proceedings of the SIGIR workshop on accessible search systems*, pages 19–26.
- P. Denis and B. Sagot. 2012. Coupling an annotated corpus and a lexicon for state-of-the-art POS tagging. *Language resources and evaluation*, 46(4):721–736.
- K. Desbiens. 2008. Les obstacles à la simplification: le cas des membres du centre d'expertise des grands organismes. *Langue, médiation et efficacité communicationnelle*. Québec.
- European Union. 2011. *How to write clearly*. Publications Office of the EU, Luxembourg.
- J. Falkenjack, E. Rennes, D. Fahlborg, V. Johansson, and A. Jönsson. 2017. Services for text simplification and analysis. In *Proceedings of the 21st Nordic Conference on Computational Linguistics*, pages 309–313.
- P.W. Foltz, W. Kintsch, and T.K. Landauer. 1998. The measurement of textual coherence with latent semantic analysis. *Discourse processes*, 25(2):285–307.
- T. François. 2011. *Les apports du traitement automatique du langage à la lisibilité du français langue étrangère*. Ph.D. thesis, Université Catholique de Louvain. Thesis Supervisors : Cédric Fairon and Anne Catherine Simon.
- T. François, L. Brouwers, H. Naets, and C. Fairon. 2014. AMesure: une formule de lisibilité pour les textes administratifs. In *Proceedings of TALN 2014*.
- T. François, A. Müller, B. Degryse, and C. Fairon. 2018. AMesure : une plateforme web d'assistance à la rédaction simple de textes administratifs. *Repères DoRiF*, 16(1).
- G. Glavaš and S. Štajner. 2015. Simplifying lexical simplification: Do we need simplified corpora? In *Proceedings of ACL-IJCNLP 2015*, pages 63–68.
- G. Gougenheim, R. Michéa, P. Rivenc, and A. Sauvageot. 1964. *L'élaboration du français fondamental (1er degré)*. Didier, Paris.
- Gouvernement du Québec. 2006. *Rédiger simplement - Principes et recommandations pour une langue administrative de qualité*. Bibliothèques et archives nationales du Québec, Québec.
- J. Kimble. 1992. Plain english: A charter for clear writing. *TM Cooley L. Rev.*, 9:1.
- J. Kimble. 1996. Writing for dollars, writing to please. *Scribes J. Leg. Writing*, 6:1.

- B. Labasse. 2001. L'institution contre l'auteur : pertinence et contraintes en rédaction professionnelle. In *Congrès annuel de l'ACPRST/CATTW, Université Laval, Québec*.
- J. Lee, W. Zhao, and W. Xie. 2016. A customizable editor for text simplification. In *Proceedings of COLING 2016: System Demonstrations*, pages 93–97.
- R. Levy and G. Andrew. 2006. Tregex and tsurgeon: tools for querying and manipulating tree data structures. In *Proceedings of LREC 2006*, pages 2231–2234.
- B. Madinier. 2009. Langage administratif et modernisation de l'état : pour une communication réussie. In *La communication avec le citoyen : efficace et accessible ? Actes du colloque de Liège 2009*, pages 55–66.
- Ministère de la Communauté française de Belgique. 2010. *Écrire pour être lu - Comment rédiger des textes administratifs faciles à comprendre ?* Ingber, Damar, Bruxelles.
- B. New, M. Brysbaert, J. Veronis, and C. Pallier. 2007. The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28(04):661–677.
- S. Nisioi, S. Štajner, S. P. Ponzetto, and L.P. Dinu. 2017. Exploring neural text simplification models. In *Proceedings of ACL2017: Short Papers*, pages 85–91.
- A. Nord. 2018. Plain language and professional writing : A research overview. Technical report, Language Council of Sweden.
- OECD. 2016. *Skills Matter : Further Results from the Survey of Adult Skills*. OECD Publishing, Paris.
- G. Paetzold and L. Specia. 2016. Unsupervised lexical simplification for non-native speakers. In *Thirtieth AAAI Conference on Artificial Intelligence*.
- G. Paetzold and L. Specia. 2017. Lexical simplification with neural ranking. In *Proceedings of EACL2017: Short Papers*, pages 34–40.
- Plain Language Action and Information Network. 2011. Federal plain language guidelines.
- J. Qiang, Y. Li, Y. Zhu, Y. Yuan, and X. Wu. 2020. Lexical simplification with pretrained encoders. In *Proceedings of AAAI 2020*, pages 8649–8656.
- H. Saggion. 2017. Automatic text simplification. *Synthesis Lectures on Human Language Technologies*, 10(1):1–137.
- C. Scarton, M. Oliveira, A. Candido Jr, C. Gasperin, and S. Aluísio. 2010. Simplifica: a tool for authoring simplified texts in brazilian portuguese guided by readability assessments. In *Proceedings of the NAACL HLT 2010: Demonstration Session*, pages 41–44.
- M. Shardlow. 2014. A survey of automated text simplification. *International Journal of Advanced Computer Science and Applications*, 4(1):58–70.
- A. Siddharthan. 2011. Text simplification using typed dependencies: A comparison of the robustness of different generation strategies. In *Proceedings of the 13th European Workshop on Natural Language Generation*, pages 2–11.
- A. Siddharthan. 2014. A survey of research on text simplification. *ITL-International Journal of Applied Linguistics*, 165(2):259–298.
- M. da Silva Conrado, A. Di Felippo, T.A.S. Pardo, and S.O. Rezende. 2014. A survey of automatic term extraction for brazilian portuguese. *Journal of the Brazilian Computer Society*, 20(1):12.
- S. Sinha, F. McDermott, G. Srinivas, and P.W.J. Houghton. 2011. Use of abbreviations by healthcare professionals: what is the way forward? *Postgraduate medical journal*, 87(1029):450–452.
- L. Specia. 2010. Translating from Complex to Simplified Sentences. In *Proceedings of Propor 2010.*, pages 30–39.
- S.M. Yimam and C. Biemann. 2018. Demonstrating par4sem-a semantic writing aid with adaptive paraphrasing. In *Proceedings of EMNLP 2018: System Demonstrations*, pages 48–53.
- X. Zhang and M. Lapata. 2017. Sentence simplification with deep reinforcement learning. In *Proceedings of EMNLP 2017*, pages 584–594.
- Z. Zhu, D. Bernhard, and I. Gurevych. 2010. A monolingual tree-based translation model for sentence simplification. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 1353–1361. Association for Computational Linguistics.
- L. Zilio, R. Wilkens, and C. Fairon. 2017. Using nlp for enhancing second language acquisition. In *Proceedings of RANLP 2017*, pages 839–846.