

AMESURE: une plateforme de lisibilité pour les textes administratifs

Thomas François¹ Laetitia Brouwers¹ Hubert Naets¹ Cédric Fairon¹
(1) CENTAL, IL&C, UCLouvain
1, Place Blaise Pascal, 1348 Louvain-la-Neuve
{thomas.francois;laetitia.brouwers;hubert.naets;cedrick.fairon}@uclouvain.be

Résumé. Cet article présente une plateforme dédiée à l'évaluation de la difficulté des textes administratifs, dans un but d'aide à la rédaction. La plateforme propose d'une part une formule de lisibilité spécialisée pour les textes administratifs, dont la conception repose sur une nouvelle méthode d'annotation. Le modèle classe correctement 58% des textes sur une échelle à 5 niveaux et ne commet d'erreurs graves que dans 9% des cas. La plateforme propose d'autre part un diagnostic plus précis des difficultés spécifiques d'un texte, sous la forme d'indicateurs numériques, mais aussi d'une localisation de ces difficultés directement dans le document.

Abstract. This paper presents a platform aiming to assess the difficulty of administrative texts, mostly for editorial assistance purposes. The platform first offers a readability formula specialized for administrative texts, the development of which required the design of a dedicated annotation procedure. The resulting model correctly classifies 58% of the texts on a 5-levels scale and commits serious misclassifications in only 9% of the cases. Moreover, the platform offers a more accurate diagnosis of the difficulty of a text in the form of numerical indicators corresponding to various textual characteristics. It also locates specific local difficulties directly in the text.

Mots-clés : formule de lisibilité, textes administratifs, aide à la rédaction.

Keywords: readability formula, administrative texts, editorial assistance.

1 Introduction

Les textes produits par les différentes administrations d'un État sont souvent perçus comme des lectures peu enthousiasmantes et particulièrement sujettes à des problèmes d'interprétation ou de compréhension, en particulier auprès de ceux que l'on peut qualifier de lecteurs faibles. D'après les dernières enquêtes PISA (De Coster *et al.*, 2011), cette catégorie engloberait presque 20% de lecteurs en moyenne en Europe. Ces problèmes de compréhension sont d'autant plus critiques dans le cas de textes administratifs que ceux-ci sont relatifs à des questions pratiques, voire parfois vitales pour un individu.

Il n'est donc pas surprenant que les administrations de divers pays aient mis en place des initiatives visant à rendre plus accessible les documents qu'elles sont amenées à produire. Plusieurs guides de rédaction simple ont ainsi vu le jour au Québec (Gouvernement du Québec, 2006), en France, en Belgique (Ministère de la Communauté française de Belgique, 2010) et au sein de l'Union européenne (Union européenne, 2011). Ils ont pour but d'orienter les personnes qui rédigent un texte administratif. Ils soulignent notamment l'importance d'écrire clairement en évitant l'ambiguïté, l'excès de détails qui finissent par occulter l'information principale, une absence de structure textuelle ou une structuration non chronologique. Ils recommandent également de mettre le lecteur au premier plan en s'adressant directement à lui, en utilisant un vocabulaire non spécialisé et en organisant l'information selon ses besoins.

Cependant, au vu des efforts répétés des administrations en ce qui concerne l'accessibilité de leurs productions, on peut douter que ces guides de rédaction simple atteignent complètement leurs objectifs. Ils nécessitent non seulement un effort important d'assimilation de la part des rédacteurs, mais également une vigilance constante, sachant qu'il est généralement assez difficile, pour un rédacteur expérimenté, de se rendre compte des difficultés qu'un lecteur non initié peut éprouver face à ses écrits. C'est pourquoi, en collaboration avec le service de la langue de la Fédération Wallonie-Bruxelles (FWB), a été monté un projet de plateforme web qui cherche à évaluer la difficulté de textes administratifs et à proposer des suggestions de simplification.

Cette plateforme vise un double objectif. D'une part, elle propose une estimation globale de la difficulté d'un texte, ce qui demande de disposer d'une formule de lisibilité spécialisée pour les textes administratifs. D'autre part, elle identifie les éléments lexicaux et syntaxiques difficiles du texte, ce qui requiert une analyse plus fine du texte, semblable à ce qui est fait dans le cadre de la simplification automatique des textes. Dans cet article, nous commençons par situer notre projet par rapport à ces deux domaines (lisibilité et simplification automatique) en mettant en évidence les défis liés au contexte particulier envisagé (les textes administratifs) (cf. section 2). Ensuite, nous présentons à la section 3 la méthodologie de conception de notre formule de lisibilité spécialisée pour les textes administratifs ainsi qu'une évaluation de celle-ci. Nous détaillons à la section 4 les éléments linguistiques identifiés comme problématiques par le système et la façon dont ceux-ci sont intégrés dans la plateforme. Nous présentons finalement à la section 5 des perspectives de développement pour cette plateforme.

2 Contexte

Depuis longtemps, la problématique de l'évaluation automatique de la difficulté des textes intéresse les chercheurs. Les premiers efforts dans ce domaine remontent au début du 20^e siècle. Ils visent à associer automatiquement des lecteurs avec des textes correspondant à leur niveau de lecture à l'aide de formules mathématiques. Ces modèles opèrent sur la base d'un nombre restreint d'informations textuelles (par ex. la longueur des mots, la longueur des phrases ou la proportion de propositions) qui sont combinées à l'aide d'algorithmes statistiques tels que la régression linéaire (par ex. chez Flesch (1948)). Plus récemment, le domaine a connu une informatisation croissante qui a conduit à la création de modèles basés sur des techniques issues de l'intelligence artificielle et qui reposent sur des variables linguistiques plus fines extraites à l'aide de techniques de traitement automatique du langage (TAL). Parmi les nombreux travaux récents, citons Collins-Thompson & Callan (2005), Heilman *et al.* (2008) ou Vajjala & Meurers (2012).

En ce qui concerne le français, les travaux sont nettement moins nombreux. La première formule de lisibilité est due à Kandel & Moles (1958) et constitue une simple adaptation de la formule de Flesch. Par la suite, Henry (1975) propose la première formule spécifique au français langue première (L1), Daoust *et al.* (1996) explorent les applications du TAL en lisibilité du français L1, tandis que François & Fairon (2013) développent une formule également basée sur le TAL pour le français langue étrangère (FLE). Cependant, aucun de ces modèles n'a été conçu spécifiquement pour les textes administratifs. En effet, créer une formule spécialisée pour un public et un type de textes particuliers requiert de disposer d'un corpus de textes dont la difficulté a été évaluée sur un échantillon de la population ciblée. En général, ce type de ressources s'obtient en collectant des textes dans des manuels scolaires organisés en fonction des niveaux d'éducation visés par le modèle. Malheureusement, pour de nombreux contextes et notamment les textes administratifs, cette solution n'est pas envisageable, ce qui freine considérablement la conception de modèles spécialisés.

Au-delà de cette difficulté méthodologique, les formules de lisibilité sont surtout adaptées à des contextes d'utilisation où il s'agit d'évaluer rapidement et globalement un nombre important de textes. Par contre, elles n'offrent pas de diagnostic précis sur les difficultés d'un texte, ce qui relève plutôt du domaine de l'aide à la rédaction ou « writability ». À l'origine, cette approche consistait à appliquer une formule de lisibilité à un texte et à réécrire celui-ci lorsque le résultat obtenu dépassait une valeur donnée. Cette approche a cependant été critiquée, car elle poussait les rédacteurs à écrire en fonction des formules, qui étaient surtout basées sur des critères de surface. Simplifier revenait donc à réduire la longueur des phrases ou des mots, ce qui ne produit pas toujours des textes plus compréhensibles (Davison & Kantor, 1982). Aujourd'hui, le recours à des technologies de TAL permet de proposer une analyse plus fine des difficultés d'un texte et divers systèmes de simplification automatisés sont apparus pour l'anglais (Chandrasekar *et al.*, 1996; Medero, 2011; Siddharthan, 2006) ou le français (Brouwers *et al.*, 2012). Les critiques qui fustigeaient l'aide à la rédaction basée sur la lisibilité semblent désormais moins pertinentes, en particulier parce que la majorité des variables utilisées dans les formules modernes sont fondées sur des connaissances psycholinguistiques. C'est dans ce contexte que nous proposons une première plateforme d'aide à la rédaction pour les textes administratifs pour le français, appelée AMESURE.

3 Une formule de lisibilité pour les textes administratifs

Comme nous l'avons expliqué, la principale difficulté de conception d'un modèle de lisibilité spécialisé pour les textes administratifs est disposer d'un corpus de textes annotés en fonction de leur difficulté de compréhension pour une population donnée. Ce type de corpus n'existant pas à notre connaissance, nous avons pris le parti d'en rassembler un à partir de documents authentiques provenant de la FWB et à destination du grand public. Il s'agit plus précisément de textes

factitifs, c'est-à-dire des documents utilitaires que le citoyen doit lire en vue de réaliser une procédure (ex. : obtenir une prime) ou pour prendre connaissance d'une problématique (ex. : comment déclarer un héritage). Nous avons ainsi collecté 88 textes relevant de huit thématiques couvertes dans les documents de la FWB (sport, culture, enfance, etc.) ainsi que 27 lettres issues de l'administration et destinées à des particuliers.

Ensuite, il a fallu annoter ces textes en fonction de leur difficulté. Pour ce faire, nous avons employé une méthode en deux étapes. Dans un premier temps, les textes ont été annotés à l'aide de la formule de Kandel & Moles (1958), ce qui a permis de les trier en fonction du score de lisibilité qui leur a été attribué (sur une échelle allant de 100 – très facile – à 0 – très difficile) et de les classer en 5 niveaux sur la base des intervalles définis sur cette échelle par de Landsheere (1963). Nous avons alors sélectionné manuellement deux textes au sein de chaque intervalle afin de disposer de documents de complexité aussi diverse que possible. Ces textes ont ensuite été lus par 10 administrés d'âges variés (de 22 à 64 ans) et titulaires, pour la plupart, d'un diplôme universitaire, via une interface d'autoprésentation segmentée¹. Celle-ci permet de mesurer le temps passé par chaque lecteur sur chaque phrase à l'aide d'un script javascript (qui évite les latences entre le client et le serveur). À la fin de la lecture, deux questions de compréhension sont posées sur le texte, afin de s'assurer qu'une lecture visant à une bonne compréhension du texte a bien été effectuée par le sujet.

Au terme de cette expérience, nous avons obtenu environ dix mesures de vitesse de lecture (ms/mot) pour chacune des phrases issues des dix textes sélectionnés. Cela nous donne 1017 observations, desquelles ont été exclues les données associées à un score inférieur à 50% au test de compréhension (62 données). Après nettoyage, nous avons calculé la vitesse moyenne de lecture des 10 sujets pour chaque phrase, en éliminant la variation spécifique aux sujets à l'aide d'un modèle à effets mixtes (Baayen *et al.*, 2008). Ensuite, une vitesse de lecture moyenne par mot a été calculée au niveau du texte. C'est cette valeur qui constitue l'indicateur de la difficulté du texte (voir Table 1). On observe une bonne corrélation ($r = 0,74$) entre le score obtenu par la formule de Kandel et Moles (score KM) et le temps moyen de lecture par mots (ms/mot) sur nos dix textes. Sur la base des temps de lecture, nous avons défini 5 niveaux de difficulté, qui constitueront l'échelle de référence pour notre formule.

Numéro du texte	Titre du texte	score KM	ms/mot	Niveau
1	La santé de votre enfant	71,3	292,8	1
2	Du couple à la famille (se séparer...)	86,5	304,9	1
3	Des chaussures... Quand les mettre aux pieds ?	81,1	315	2
4	A l'école d'une alimentation saine	75,8	324,4	2
5	L'enseignement spécialisé	46,2	339,7	3
6	Lettre pour la semaine européenne de la vaccination	40,6	340,5	3
7	Cumuls de pensions	57,5	372,3	4
8	Liquidation des subventions ordinaires 2004	15	376,6	4
9	Déclaration de succession	57	379	5
10	Tax shelter	36,5	390	5

TABLE 1 – Classement des textes selon leur temps de lecture moyen.

Pour la seconde étape de l'annotation, nous avons fait appel à des experts issus de l'administration de la FWB. Il a été demandé à chacun d'eux de classer 15 textes parmi les 105 restant à annoter, en fonction des 5 niveaux de difficulté définis lors de la première étape. Pour les assister dans cette tâche, un guide d'annotation leur a été fourni. Il inclut, comme exemples représentatifs de chaque niveau, les textes de la Table 1 numérotés 1, 3, 6, 7 et 10. Au terme de cette annotation, nous avons obtenu 267 avis sur 105 textes, soit une moyenne de 2,5 avis par textes. L'accord inter-annotateurs n'est pas très élevé : l'alpha de Krippendorff (1980) moyen sur les 7 batchs de 15 textes ne dépasse pas 0,37. Notons toutefois que ce type de tâche d'annotation est particulièrement difficile, comme le montre le kappa de 0,398 obtenu pour une tâche d'annotation de la difficulté de mots dans SemEval 2012 (Specia *et al.*, 2012). Au terme de ce processus d'annotation, chaque texte s'est vu attribuer le niveau moyen des annotations des experts, réduisant les variations personnelles.

Une fois le corpus d'entraînement annoté, la seconde étape de conception de notre formule de lisibilité a consisté à identifier et paramétrer un ensemble de variables linguistiques qui sont liées avec la difficulté des textes, par un lien causal ou simplement corrélational. Nous avons ainsi pris en considération 344 variables qui se répartissent en trois grandes classes : lexicale (fréquence lexicale, diversité lexicale, types de voisins orthographiques et longueur des mots), syntaxique (longueur des phrases, formes verbales, ratios de catégories de discours) et sémantique (niveau de personnalisation du texte, densité des idées et cohésion du texte). Ces variables ont été décrites en détail dans François & Fairon (2013).

1. Celle-ci est disponible à l'adresse suivante : <http://amesure-tests.cental.be/>

Enfin, pour entraîner le modèle, nous avons opéré une sélection de variables à l'aide d'un double critère. Tout d'abord, nos variables ont été ordonnées en fonction de leur capacité prédictive, comme le recommandent Guyon & Elisseeff (2003). Nous avons calculé une corrélation de Spearman entre chacune des variables et le niveau de difficulté des textes. Dans un second temps, la meilleure variable au sein de chacune des sous-familles a été retenue. Le but de cette procédure était à la fois de maximiser la quantité d'information à disposition du modèle tout en limitant le nombre de variables reprises dans le modèle, étant donné qu'il doit s'intégrer dans une architecture en temps réel. Au terme de cette étape de sélection, 10 variables ont été retenues. Elles sont présentées à la Table 2 avec la corrélation obtenue pour chacune d'elles. On remarque en particulier que la complexité syntaxique (NMP et CON_PRO) importe davantage que les indices lexicaux dans les textes administratifs, à l'inverse de ce qui a été observé pour les textes de FLE (François & Fairon, 2013).

Nom	Description de la variable	corr.
ML3	Modèle unigramme lissé basé sur les fréquences de Lexique3 (New <i>et al.</i> , 2001)	-0,32
MedianFFFDV	Médiane des fréquences des verbes du texte	-0,47
PAGoug_8000	Proportion de mots absents des 8000 premiers mots de la liste longue de Gougenheim	0,44
TTR_W	Type-Token ratio calculé sur les lemmes	-0,21
PM8	Proportion de mots de plus de 8 lettres	0,40
mean_freqCumNeigh	Moyenne de la distribution des fréquences cumulées des voisins orthographiques	0,50
NMP	Nombre moyen de mots par phrase	0,64
PPasse_C	Proportion de verbes conjugués au participe passé	0,46
CON_PRO	Rapport du nombre de conjonctions sur celui des pronoms	0,54
PP1P2	Nombre de pron. pers. S1 et S2 / mots	-0,42

TABLE 2 – Liste des variables sélectionnées dans le modèle avec leur corrélation de Spearman.

Enfin, un modèle capable d'estimer la difficulté des textes administratifs a été entraîné à l'aide d'une machine à vecteurs de support (SVM) dans laquelle les coefficients ont été régularisés via une norme L2. Une première série d'explorations a permis de sélectionner un *kernel* linéaire, pour lequel le coût (C) a été optimisé par *grid-search*. L'efficacité de ces modèles a été estimée à l'aide d'une procédure de validation croisée à 10 échantillons. Le modèle retenu ($C = 0.05$) atteint une exactitude de 58% pour la classification et une exactitude contiguë² de 91%. Il s'agit donc d'un gain de 38% en exactitude et de 39% en exactitude contiguë par rapport au hasard. Pour comparaison, François & Fairon (2013) obtiennent respectivement 50% en exactitude et 80% en exactitude contiguë pour leur modèle sur des textes de FLE. Cela semble indiquer que le modèle se comporte plutôt bien malgré le nombre réduit de données d'entraînement. On notera toutefois qu'il a tendance à rabattre les textes de niveau 1 au niveau 2 et ceux de niveau 5 en 4, à cause du nombre plus réduit de données pour ces catégories.

4 Plateforme d'aide à la rédaction

Cette formule de lisibilité pour les textes administratifs a été intégrée au sein d'une plateforme plus globale d'aide à la rédaction³. En plus de cette estimation globale de la difficulté du texte soumis, la plateforme propose actuellement deux types de diagnostics. Le premier fournit une analyse plus détaillée de la complexité d'un texte sur le modèle de Coh-Matrix (Graesser *et al.*, 2004). Il offre ainsi à l'utilisateur une série d'indicateurs mesurant la complexité des différentes dimensions d'un document administratif. Pour l'instant cinq indicateurs ont été retenus : PAGoug_8000, NMP, CON_PRO et PP1P2 (voir Table 2), ainsi qu'une mesure de la cohérence moyenne du texte. Cette dernière est estimée comme le cosinus moyen de toutes les paires de phrases adjacentes du texte représentées sous la forme d'un vecteur de mots dans un espace de dimensions réduites (à l'aide d'une analyse sémantique latente).

Le second diagnostic va plus loin en cherchant à identifier précisément des points d'achoppement probables dans le texte, qu'il s'agisse d'un mot rare ou d'une structure syntaxique considérée comme plus difficile à analyser. C'est à ce niveau que la plateforme rejoint les travaux en simplification de textes, puisque notre approche adopte une approche semblable à ceci près qu'elle s'arrête au niveau de l'identification des éléments à simplifier sans effectuer automatiquement de simplifications. À ce stade de développement de la plateforme, ce module se limite à repérer les mots rares sur la base de leur fréquence, ainsi qu'à identifier trois types de constructions qui peuvent poser des difficultés de lecture. La Figure 1 propose un exemple d'analyse sur un texte court et présente l'interface de la plateforme. Les mots rares sont grasseyés en rouge et les structures complexes soulignées dans une couleur qui dépend de leur type.

2. Il s'agit de la proportion de prédictions qui s'éloignent au plus d'un niveau par rapport au niveau de référence. Son emploi, commun en lisibilité, se justifie par la difficulté qu'ont les experts humains eux-mêmes à accorder leurs jugements.

3. La plateforme est gratuitement accessible à l'adresse suivante : <http://cental.uclouvain.be/amesure/>.

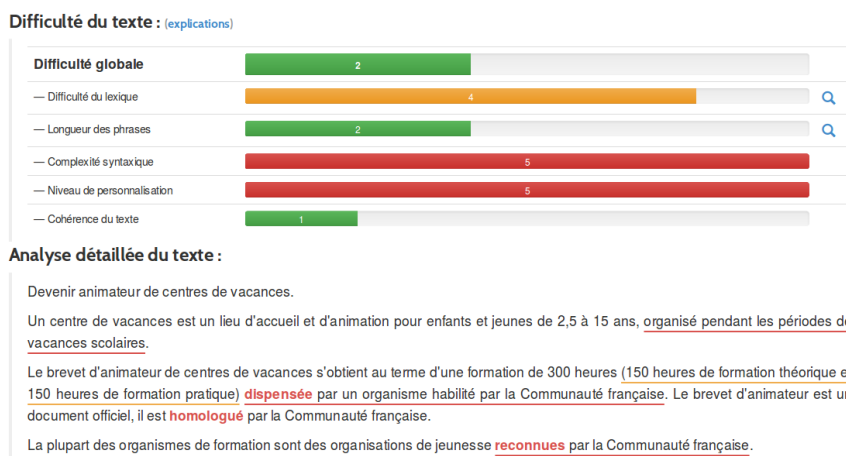


FIGURE 1 – Capture d'écran des résultats de la plateforme AMESURE sur un texte.

La détection des mots rares est réalisée en comparant les logarithmes des fréquences des formes fléchies du texte (tirées de Lexique3 (New *et al.*, 2001)) à un seuil de référence (−14 dans notre exemple), qui pourrait cependant être manipulé par l'utilisateur en fonction du public qu'il vise. Quant aux trois constructions syntaxiques détectées, elles regroupent (1) le passif ; (2) les éléments entre parenthèses et (3) les propositions subordonnées. En effet, lorsque le verbe principal est à la voix passive, la phrase ne suit pas l'ordre SVO auquel s'attend le lecteur, ce qui peut poser davantage de problèmes de compréhension. De même, les éléments entre parenthèses, s'ils sont accumulés, ont tendance à occulter l'information principale qui se retrouve noyée au milieu des informations secondaires. Enfin, les propositions subordonnées contribuent également à allonger et alourdir la phrase, la rendant moins claire. C'est pourquoi de nombreux manuels de rédaction simple (Gouvernement du Québec, 2006; Ministère de la Communauté française de Belgique, 2010; Union européenne, 2011) recommandent de simplifier ce type de structure.

5 Conclusion et perspectives

Cet article a présenté AMESURE, une plateforme dédiée à l'évaluation de la difficulté des textes administratifs, dont le but principal est d'aider les rédacteurs de ce type de textes à produire des documents plus accessibles. La principale contribution de cet article consiste en une formule de lisibilité spécialisée pour les textes administratifs, dont les performances sont comparables à celles d'autres modèles actuels pour le français, malgré un nombre réduit de textes pour l'entraînement. Cela a été rendu possible grâce à une technique d'annotation innovante dans le domaine de la lisibilité qui se base sur la mesure de la vitesse de lecture moyenne d'un groupe de lecteurs. À l'aide de celle-ci, nous avons défini une échelle de difficulté et un guide d'annotation ayant ensuite servi à un groupe d'experts à annoter le reste du corpus. Cette technique pourrait être réutilisée pour entraîner d'autres modèles de lisibilité spécifiques lorsque, comme dans notre cas, il n'existe pas de textes gradués disponibles.

Par ailleurs, cette plateforme intègre divers indicateurs des difficultés présentes dans les textes administratifs relevant de plusieurs dimensions : complexité lexicale du texte, complexité de ses structures syntaxiques, taux de cohérence, etc. Par ailleurs, un diagnostic est également fourni concernant un certain nombre de points d'achoppement relatifs à la rareté du vocabulaire employé et à des structures syntaxiques considérées comme potentiellement problématiques.

À notre connaissance, cette plateforme constitue le premier outil librement accessible pour le français qui offre à la fois une évaluation de la lisibilité d'un texte et un diagnostic plus précis. Il s'agit cependant d'une première étape et les diagnostics, en particulier, pourraient être étoffés. Tout d'abord, on peut imaginer d'autres méthodes pour repérer les mots difficiles, soit en se focalisant sur les termes techniques à l'aide d'outils d'extraction terminologiques, soit en évaluant la difficulté d'un mot à partir de plusieurs critères (et pas seulement de la fréquence) à l'instar de ce qui est proposé par Gala *et al.* (2013). En ce qui concerne les structures syntaxiques complexes, nous comptons également repérer davantage de structures problématiques, telles que les phrases négatives et à double négation, les interrogatives indirectes, les phrases qui ne respectent pas l'ordre sujet-verbe-objet (outre la forme passive), les clivées, les compléments circonstanciels, etc.

Références

- BAAYEN R. H., DAVIDSON D. J. & BATES D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, **59**(4), 390–412.
- BROUWERS L., BERNHARD D., LIGOZAT A.-L. & FRANÇOIS T. (2012). Simplification syntaxique de phrases pour le français. In *Actes de la Conférence Conjointe JEP-TALN-RECITAL*, p. 211–224.
- CHANDRASEKAR R., DORAN C. & SRINIVAS B. (1996). Motivations and methods for text simplification. In *Proceedings of the 16th conference on Computational Linguistics*, volume 2, p. 1041–1044.
- COLLINS-THOMPSON K. & CALLAN J. (2005). Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology*, **56**(13), 1448–1462.
- DAOUST F., LAROCHE L. & OUELLET L. (1996). SATO-CALIBRAGE : 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.
- DAVISON A. & KANTOR R. (1982). On the failure of readability formulas to define readable texts : A case study from adaptations. *Reading Research Quarterly*, **17**(2), 187–209.
- DE COSTER I., BAIDAK N., MOTIEJUNAITE A. & NOORANI S. (2011). *Teaching Reading in Europe : Contexts, Policies and Practices*. Rapport interne, Education, Audiovisual and Culture Executive Agency, European Commission.
- DE LANDSHEERE G. (1963). Pour une application des tests de lisibilité de Flesch à la langue française. *Le Travail Humain*, **26**, 141–154.
- FLESCH R. (1948). A new readability yardstick. *Journal of Applied Psychology*, **32**(3), 221–233.
- FRANÇOIS T. & FAIRON C. (2013). Les apports du TAL à la lisibilité du français langue étrangère. *Traitement Automatique des Langues (TAL)*, **54**(1), 171–202.
- GALA N., FRANÇOIS T. & FAIRON C. (2013). Towards a french lexicon with difficulty measures : Nlp helping to bridge the gap between traditional dictionaries and specialized lexicons. In *Electronic lexicography in the 21st century : thinking outside the paper (eLex2013)*, p. 132–151.
- GOVERNEMENT DU QUÉBEC (2006). *Rédiger simplement - Principes et recommandations pour une langue administrative de qualité*. Québec : Bibliothèques et archives nationales du Québec.
- GRAESSER A., MCNAMARA D., LOUWERSE M. & CAI Z. (2004). Coh-Metrix : Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, **36**(2), 193–202.
- GUYON I. & ELISSEEFF A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, **3**, 1157–1182.
- HEILMAN M., COLLINS-THOMPSON K. & ESKENAZI M. (2008). An analysis of statistical models and features for reading difficulty prediction. In *Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications*, p. 1–8.
- HENRY G. (1975). *Comment mesurer la lisibilité*. Bruxelles : Labor.
- KANDEL L. & MOLES A. (1958). Application de l'indice de Flesch à la langue française. *Cahiers Études de Radio-Télévision*, **19**, 253–274.
- KRIPPENDORFF K. (1980). *Content analysis : An introduction to its methodology*. Beverly Hills, CA : Sage.
- MEDERO, J. AND OSTENDORF M. (2011). Identifying targets for syntactic simplification. In *Proceedings of the SLaTE 2011 workshop*.
- MINISTÈRE DE LA COMMUNAUTÉ FRANÇAISE DE BELGIQUE (2010). *Écrire pour être lu - Comment rédiger des textes administratifs faciles à comprendre ?* Bruxelles : Ingber, Damar.
- NEW B., PALLIER C., FERRAND L. & MATOS R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L'Année Psychologique*, **101**, 447–462.
- SIDDHARTHAN A. (2006). Syntactic simplification and text cohesion. *Research on Language and Computation*, **4**(1), 77–109.
- SPECIA L., JAUHAR S. K. & MIHALCEA R. (2012). Semeval-2012 task 1 : English lexical simplification. In *Proceedings of the Sixth International Workshop on Semantic Evaluation*, p. 347–355.
- UNION EUROPÉENNE (2011). *Rédiger clairement*. Luxembourg : Office des publications de l'Union européenne.
- VAJJALA S. & MEURERS D. (2012). On improving the accuracy of readability classification using insights from second language acquisition. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, p. 163–173.