Simplification automatique de texte dans un contexte de faibles ressources

Sadaf Abdul Rauf, Anne-Laure Ligozat, François Yvon, Gabriel Illouz, Thierry Hamon

Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 31e édition),
Traitement Automatique des Langues Naturelles (TALN, 27e édition),
Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition)
Nancy, France, 08-19 juin 2020

Simplification automatique de texte dans un contexte de faibles ressources

Sadaf Abdul Rauf[1], Anne-Laure Ligozat[1, 2], Francois Yvon[1],
Gabriel Illouz[1] and Thierry Hamon[1, 3]
(1) Université Paris-Saclay, CNRS, LIMSI, 91400, Orsay, France
(2) Université Paris-Saclay, CNRS, ENSIIE, LIMSI 91400, Orsay, France
(3) Université Sorbonne Paris Nord, 93430, Villetaneuse, France
{firstName.lastName}@limsi.fr

RÉSUMÉ

La simplification de textes a émergé comme un sous-domaine actif du traitement automatique des langues, du fait des problèmes pratiques et théoriques qu'elle permet d'aborder, ainsi que de ses nombreuses applications pratiques. Des corpus de simplification sont nécessaires pour entrainer des systèmes de simplification automatique ; ces ressources sont toutefois rares et n'existent que pour un petit nombre de langues. Nous montrons ici que dans un contexte où les ressources pour la simplification sont rares, il reste néanmoins possible de construire des systèmes de simplification, en ayant recours à des corpus synthétiques, par exemple obtenus par traduction automatique, et nous évaluons diverses manières de les constituer.

ABSTRACT

Automatic Text Simplification : Approaching the Problem in Low Resource Settings for French

Sentence simplification has emerged as an active area of research in recent years owing to its utility in natural language processing as well as human learning studies. Simplification corpora are required to build such automatic simplification systems. We show that where resources for simplification are scarce, it is still possible to build simplification systems. We show the effectiveness of translations as a synthetic corpus for the simplification task and present an analysis of the two metrics often used to evaluate simplification, i.e. BLEU and SARI, in terms of their ability to predict complexity level of a text.

MOTS-CLÉS : Simplification de textes, compression de texte, corpus synthétique, apprentissage par transfert cross-lingue.

KEYWORDS: Text Simplification, Sentence Compression, Synthetic Corpus, Cross-Lingual Transfert Learning.

1 Introduction

La lecture et la compréhension constituent une contribution majeure à la courbe d'apprentissage d'une langue pour un apprenant. La complexité d'une phrase empêche souvent la bonne compréhension et constitue un obstacle majeur dans la chaîne d'apprentissage. Ceci s'applique spécifiquement à certains groupes de personnes, par exemple les enfants (De Belder & Moens, 2010), les personnes ayant des difficultés d'apprentissage (Rello *et al.*, 2013a; Huenerfauth *et al.*, 2009; Fajardo *et al.*, 2013) ou des difficultés spécifiques comme des formes de dyslexie (Rello *et al.*, 2013b; McCarthy & Swierenga,

Volume 2 : Traitement Automatique des Langues Naturelles, pages 332–341.
HAL : hal-02784783.

2010) ou d'aphasie (Canning & Tait, 1999), ou encore des troubles du spectre autistique (Evans *et al.*, 2014; Barbu *et al.*, 2015). Disposer de versions simplifiées d'un texte peut grandement en faciliter la lisibilité et la compréhension pour ces populations, et pourrait aussi bénéficier à des apprenants d'une langue étrangère.

La simplification automatique de texte (SAT) construit une version plus simple de textes complexes afin qu'ils soient plus facilement compréhensibles et intelligibles. La simplification implique des transformations élaborées du texte d'origine telles que le fractionnement, la suppression et la paraphrase. Nous nous limitons ici à des transformations qui s'appliquent à des phrases isolées.

La plupart des approches de simplification automatique de texte considèrent le processus de simplification comme une tâche de traduction monolingue, l'algorithme de traduction apprenant la simplification en réécrivant à partir du corpus parallèle simple-complexe (Daelemans *et al.*, 2004; Zhu *et al.*, 2010; Zhang & Lapata, 2017; Nisioi *et al.*, 2017).

Pour qu'un algorithme d'apprentissage automatique puisse apprendre ces transformations, il faut lui fournir suffisamment d'exemples de simplifications appariant des couples de phrases simples et complexes. La qualité d'un système de simplification automatique de texte dépendra fortement de la qualité et de la quantité des corpus de simplification ainsi que la qualité de l'algorithme d'apprentissage.

Il existe de nombreuses ressources linguistiques pour le français, mais en ce qui concerne la tâche de simplification, il s'agit d'une langue peu dotée. Nous ne connaissons aucun corpus de grande taille disponible gratuitement. Nous avons collecté des corpus de diverses sources (Grabar & Cardon, 2018; Brouwers *et al.*, 2014), mais ils restent trop petits pour développer des systèmes de simplification basés sur des méthodes d'apprentissage ; nous les utilisons uniquement comme corpus de test.

Nous proposons d'utiliser des traductions automatiques pour composer un corpus parallèle synthétique. Les corpus synthétiques sont souvent utilisés en traduction automatique (Lambert *et al.*, 2011; Abdul Rauf *et al.*, 2016; Sennrich *et al.*, 2016; Burlot & Yvon, 2018) mais leur utilisation reste peu explorée pour la simplification automatique. (Aprosio *et al.*, 2019) exploitent des données synthétiques pour créer des systèmes de simplification, mais en s'appuyant sur des données de simplification de "gold standard" en italien qui leur permettent d'entraîner le système "complexificateur", qu'ils utilisent pour traduire les phrases simples en complexes. Dans notre étude, nous faisons l'hypothèse qu'aucune donnée n'est initialement disponible et nous étudions s'il est possible de construire un système de simplification raisonnable à partir uniquement de traductions automatiques de phrases complexes et simples. À cet égard, nous présentons la première tentative d'utilisation d'un corpus synthétique pour aborder le problème de la simplification automatique des phrases et montrons qu'il s'agit d'une approche viable pour réaliser des systèmes de simplification dans des contextes à faibles ressources.

2 État de l'art

La plupart des approches de simplification automatique de textes considèrent le problème comme une tâche de traduction monolingue, l'algorithme de traduction apprenant la réécriture de la simplification à partir du corpus parallèle complexe-simple (Daelemans *et al.*, 2004; Zhu *et al.*, 2010; Zhang & Lapata, 2017; Nisioi *et al.*, 2017). Considérant la simplification comme un problème d'apprentissage automatique, la traduction monolingue a été opérée en utilisant toutes les techniques de traduction automatique, y compris la traduction à base de syntaxe (Zhu *et al.*, 2010), la traduction à base

de segments (Wubben *et al.*, 2012), l'hybridation du modèle de simplification avec la traduction automatique (Narayan & Gardent, 2014) et plus récemment diverses architectures neuronales (Zhang & Lapata, 2017; Nisioi *et al.*, 2017; Zhao *et al.*, 2018; Korhonen *et al.*, 2019; Surya *et al.*, 2019). Cependant, aucune de ces études n'a abordé la simplification dans un scénario de ressources limitées. Une exception est (Aprosio *et al.*, 2019) qui sélectionne des phrases simples à partir de corpus monolingues et produit les phrases complexes correspondantes en utilisant un "complexificateur" entrainé sur le corpus de simplification italien gold standard.

La qualité du corpus d'entraînement est un facteur important et a été largement discutée, en particulier pour le corpus *Wikipedia simple* (Xu *et al.*, 2015; Scarton *et al.*, 2018). Newsela (Xu *et al.*, 2015) est devenu une des principales ressources en la matière, car il propose des simplifications manuelles à plusieurs niveaux, du niveau 0 comprenant le texte original au niveau 4 comprenant le texte le plus simplifié ; cette propriété reste cependant sous-exploitée et aucun des travaux de l'état de l'art ne donne d'analyse approfondie des niveaux de simplification les plus utiles. Des travaux antérieurs comme (Alva-Manchego *et al.*, 2017) et (Scarton *et al.*, 2018) ont pris en compte ces niveaux et les ont utilisés en exploitant les niveaux de simplification adjacents (par exemple 0-1, 1-2), mais n'ont pas étudié les différences entre les niveaux. (Zhang & Lapata, 2017) ont utilisé des niveaux non-adjacents (par exemple 0-2, 1-4), tandis que (Scarton & Specia, 2018) ont utilisé les différents niveaux de lisibilité pour construire des systèmes de simplification pour différents publics. Dans cet article, nous présentons des systèmes de simplification et des analyses avec toutes les paires de niveaux et étudions également l'effet combiné de différents niveaux sur la simplification.

Un autre objectif de notre travail est d'étudier si les résultats obtenus pour l'anglais peuvent être reproduits pour d'autres langues. Pour le français, des approches basées sur des règles existent (Brouwers *et al.*, 2014) mais aucun système basé sur l'apprentissage automatique ne peut être construit en raison du manque de corpus de simplification. Un troisième objectif est d'utiliser des données synthétiques, comme cela a été fait pour la traduction automatique dans (Lambert *et al.*, 2011; Abdul Rauf *et al.*, 2016; Sennrich *et al.*, 2016; Burlot & Yvon, 2018) et d'étudier les niveaux de simplification et l'impact des corpus synthétiques.

3 Corpus de simplification synthétique

Nous avons construit des corpus de simplification français synthétiques en traduisant deux corpus de simplification anglais avec l'outil Google translate (version de mai 2019) [1]. Pour obtenir une mesure de la qualité des traductions, nous avons traduit le corpus de test anglais-français WMT14 [2] dans les mêmes conditions et avons obtenu un score BLEU (Papineni *et al.*, 2002) de $35,6$ sur l'ensemble de test tokenisé, comparable au meilleur système de cette évaluation, mais significativement moins bon que l'état de l'art actuel de la traduction automatique anglais-français.

Le corpus anglais qui nous sert de point de départ est Newsela (Xu *et al.*, 2015), qui a été créé pour fournir du matériel de lecture destiné à l'enseignement pré-universitaire. Chaque article a été réécrit quatre fois pour des enfants de différents niveaux. Le niveau 0 correspond à l'article original, et le même article apparaît pour 4 niveaux, du niveau 1 au niveau 4, le niveau 4 étant le plus simple. Nous avons couplé le corpus dans toutes les configurations possibles de $complex : simple$, où $complex$ comprend tous les textes d'un niveau de complexité donné l et simple comprend les textes ayant un niveau de complexité inférieur à l. Les corpus ainsi construits associent le niveau 0 aux niveaux

1. https://translate.google.com
2. https://www.statmt.org/wmt14/translation-task.html

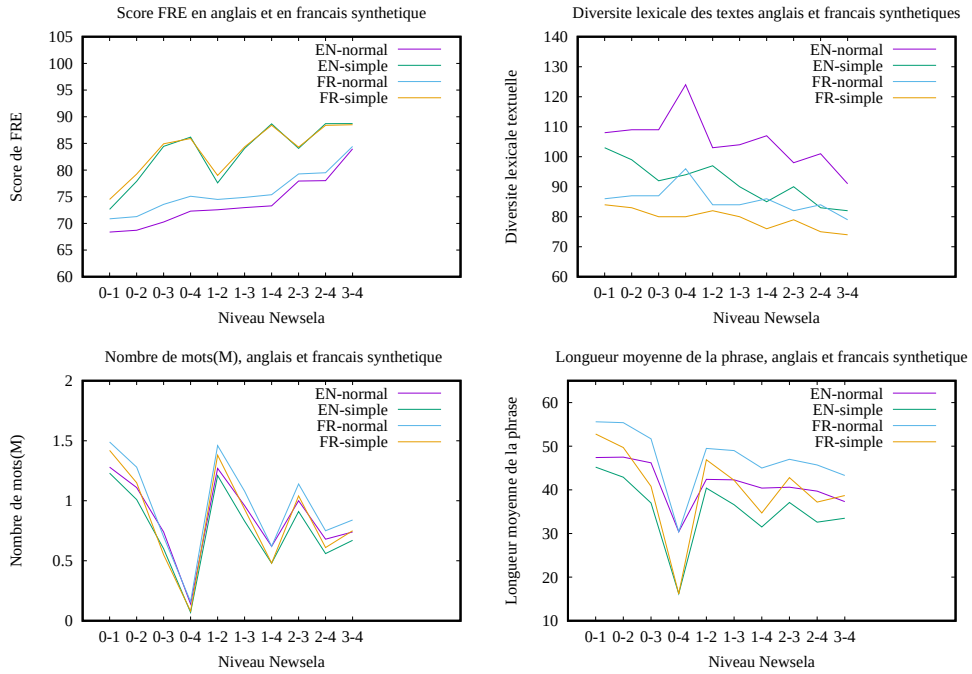

FIGURE 1 – Propriétés lexicales et de complexité des corpus originaux et traduits

$\{1, 2, 3, 4\}$, 1 à $\{2, 3, 4\}$, 2 à $\{3, 4\}$ et 3 à 4 comme indiqué dans le tableau 2.

La figure 1 présente l'analyse lexicale et la complexité des textes originaux et traduits. Nous constatons que les textes conservent leur niveau de lisibilité après traduction, comme le montrent les paramètres. Le comportement suivant est manifesté par les corpus de simplification anglais et français synthétique, où 0 est la phrase complexe et 4 le niveau le plus simple :

Niveau de complexité : Nous utilisons le score Flesch Reading Ease (Flesch, 1948) pour l'anglais et son adaptation en français proposée par (Kandel & Moles, 1958), implantés respectivement par les équations (1) et (2) ci-dessous. Ces scores varient entre 0 et 100. Une valeur plus élevée signifie que le texte est facile à lire et une valeur plus faible signifie que la difficulté est plus élevée.

$$FRE(Anglais) = 206.835 - 1.015(\frac{\frac{total}{words}}{\frac{total}{sentences}} - 84.6(\frac{\frac{total}{syllables}}{\frac{total}{words}}) \quad (1)$$

$$FRE(Francais) = 207 - 1.015(\frac{\frac{total}{words}}{\frac{total}{sentences}} - 73.6(\frac{\frac{total}{syllables}}{\frac{total}{words}}) \quad (2)$$

Comme on le voit sur la figure 1 en haut à gauche, après traduction, le corpus simple français synthétique se situe au même niveau de lisibilité que l'anglais simple original. Il en est de même pour les phrases complexes. On peut donc conclure qu'au moins en surface, les textes traduits conservent la classe de complexité des textes originaux.

Diversité lexicale textuelle : La diversité lexicale (DL) [3] est une mesure fréquemment utilisée pour évaluer la niveau de complexité d'un texte, un indice de DL est élevé indiquant qu'il est sera plus difficile à lire. Le score DL identifie clairement la classe de complexité du corpus comme le montre le sous-graphique en haut à droite (figure 1). On observe que du côté complexe, le score de diversité

3. Les mesures sont calculées avec https://pypi.org/project/lexicalrichness/

lexicale des textes augmente, alors que du côté simple, il diminue d'un niveau à l'autre, ainsi pour le côté complexe (1-2 = 103, 1-3 = 104, 1-4 = 107) et le côté simple (1-2 = 97, 1-3= 90, 1-4 = 85).

Longueur moyenne de la phrase et nombre de mots : Le nombre de mots diminue à mesure que le niveau de simplicité augmente. Ce comportement s'applique aussi bien aux mots complexes qu'à leur équivalent simple parmi toutes les paires de niveaux. La longueur moyenne des phrases est également beaucoup plus faible pour les textes simples. Ceci est illustré graphiquement dans les deux derniers sous-graphes.

On note que les phrases parallèles peuvent être très différentes dans chaque paire de niveaux, car elles sont sélectionnées par l'aligneur de phrases (Moore, 2002) en fonction du score de l'aligneur.

4 Évaluation

4.1 Protocole d'évaluation

Nous évaluons le système de simplification en utilisant les scores BLEU et SARI. BLEU (K. Papineni & Ward, 1998) est une métrique de traduction automatique qui s'appuie sur une comparaison de surface entre la sortie du système et la phrase de référence, en utilisant des correspondances de n-grammes et une pénalité de brièveté. Il donne une mesure de *l'adéquation*. SARI (Xu *et al.*, 2015) compare la phrase de sortie avec la phrase d'entrée ainsi qu'avec les phrases de référence. Il récompense les fragments qui sont validés par une des références, alors qu'ils n'apparaissaient pas dans l'entrée. BLEU est bien corrélé aux scores humains pour l'évaluation de la grammaticalité et, dans une moindre mesure, du transfert sémantique, tandis que SARI est fortement corrélé aux scores humains pour l'évaluation de la simplicité.

Pour illustrer le fonctionnement des deux métriques, le tableau 1 présente un exemple et des scores obtenus à partir de (Xu *et al.*, 2015). Pour cet exemple, BLEU attribue le même score à OUTPUT-2 et OUTPUT-3 puisque les deux mots "now" et "currently" apparaissent dans les phrases de référence, SARI en revanche donne un meilleur score à OUTPUT-2 sur la base de l'utilisation du mot "now", qui n'était pas présent dans la phrase originale.

INPUT	About 95 species are currently accepted .		
REF-1	About 95 species are currently known .		
REF-2	About 95 species are now accepted .		
REF-3	95 species are now accepted .		
	System Output	**B**LEU	**S**ARI
OUTPUT-1	About 95 you now get in .	0.1562	0.2683
OUTPUT-2	About 95 species are now agreed .	0.6435	0.7594
OUTPUT-3	About 95 species are currently agreed.	0.6435	0.5890

TABLE 1 – BLEU par rapport à SARI : exemple de score de (Xu *et al.*, 2015).

Pour l'évaluation du système, deux corpus de tests ont été utilisés pour l'anglais et le français : le corpus *self-test*, composé de 10% de phrases du système en cours de construction (non inclus dans l'entraînement) ; et des corpus de test standard construits par des humains. Pour l'anglais, ce corpus est le corpus Turk (Xu *et al.*, 2015) ayant 8 phrases de référence et, pour le français, le corpus de simplification (Grabar & Cardon, 2018) a été utilisé (Fr-test). Turk est basé sur la Wikipédia simple en anglais et Fr-test est basé sur la Wikipédia française et sur Vikidia.

	English						Synthetic French					
	nombre de	*Self-test*			*Turk-test*		*nombre de*	*Self-test*			*Fr-test*	
Niveau	*phrases*	**BLEU**	**SARI**	**FRE**	**BLEU**	**SARI**	*phrases*	**BLEU**	**SARI**	**FRE**	**BLEU**	**SARI**
0-1	27047	63.66	33.51	71.70	60.51	32.74	26809	49.28	31.50	74.62	6.64	34.96
0-2	23556	43.81	33.40	78.77	48.20	29.02	23175	8.20	18.75	82.19	3.41	34.03
0-3	16114	22.52	31.77	87.92	21.84	21.82	13483	8.33	24.96	87.35	2.96	33.73
0-4	4644	4.67	34.25	86.23	2.13	13.26	4587	4.01	33.91	83.90	1.13	31.50
1-2	29973	55.30	33.96	77.47	51.67	30.45	29613	8.86	18.82	87.73	3.14	33.73
1-3	22728	30.55	34.28	85.03	33.88	26.06	22103	8.49	24.05	88.17	3.82	34.01
1-4	15292	17.26	33.71	93.78	13.56	19.28	13899	11.50	32.49	94.52	2.36	33.98
2-3	24668	40.58	33.74	84.36	43.40	28.24	24238	35.50	33.14	86.29	4.75	34.68
All	226351	50.10	37.69	85.72	47.10	26.22	218079	43.46	36.88	86.37	4.27	34.18

TABLE 2 – Scores BLEU et SARI pour l'anglais et le français traduit, obtenus en faisant varier les niveaux du corpus Newsela utilisés à l'apprentissage.

4.2 Évaluation expérimentale

Notre cadre expérimental a été conçu pour répondre aux questions suivantes :

— En l'absence de ressources suffisantes pour la simplification, est-il possible de construire un système de simplification automatique « raisonnable » pour une langue ?
— Quelle est l'efficacité de l'utilisation de corpus synthétiques pour la tâche de simplification ?

Nous avons utilisé OpenNMT-py (Klein *et al.*, 2017) pour développer des modèles de simplification en anglais et en français. Un réseau neuronal récurrent (RNN) seq2seq à deux couches a été utilisé avec 500 unités de mémoire à long et court terme (LSTM) dans chaque couche. L'apprentissage a été optimisé via l'optimiseur Adam avec un taux d'apprentissage de 0,001 pour toutes les expériences. Nous avons utilisé une taille de lot de 64 et une taille de lot de 128 pour les corpus plus grands. Une validation est effectuée tous les 5000 pas d'apprentissage.

5 Discussion et conclusion

Le tableau 2 présente les résultats obtenus pour tous les systèmes de simplification pour l'anglais et le français et pour tous les niveaux de simplification. Nous utilisons le score FRE comme mesure supplémentaire pour évaluer la simplicité de la production. Les scores FRE ne sont calculés que pour l'"auto-test", car il s'agit d'un ensemble de tests communs à l'anglais et au français qui nous permet d'établir une comparaison entre les systèmes.

On observe que les résultats sont très variés. Pour les systèmes de simplification construits à partir de l'anglais Newsela, une première observation est que les scores BLEU baissent de manière drastique lorsque l'écart entre niveaux de simplicité augmente. Par exemple, pour le self test pour 0 $\{0-1 \Rightarrow 63.66, 0-2 \Rightarrow 43.81, 0-3 \Rightarrow 22.52, 0-4 \Rightarrow 4.67\}$ et ensuite un bond pour $1-2 \Rightarrow 55.30$ qui baisse encore pour $1-3 \Rightarrow 30.55$ et encore pour $1-4 \Rightarrow 17.26$ et enfin un bon score pour $2-3 \Rightarrow 40.58$ mais inférieur à $1-2$. Les scores BLEU supérieurs à 40 sont principalement obtenus pour des niveaux de simplicité consécutifs, c'est-à-dire $0-1$, $0-2$, $1-2$, $1-3$. Cette tendance se retrouve à

la fois dans les corpus Self et Turk.

Cependant, ce phénomène n'est pas aussi unanimement démontré par les systèmes construits à partir de la Newsela française synthétique. Cette tendance est illustrée par les résultats de l'évaluation des systèmes utilisant le test français, à l'exception du score pour $1-3$, et par self-test, à l'exception du score pour $1-2$. Ici aussi, les bons systèmes sur self-test sont les systèmes basés sur des niveaux consécutifs de simplicité, par exemple $0-1$, $1-2$, $1-3$.

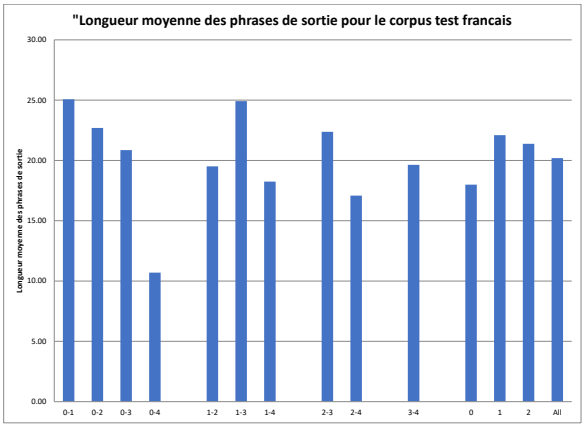

FIGURE 2 – Les longueurs moyennes des phrases de sortie pour le test de français sur différents niveaux de complexité sur Newsela.

SARI, en revanche, donne des résultats relativement stables à tous les niveaux de simplification pour le corpus self-test en anglais et en français. La même tendance à la baisse du score par niveau de simplicité est observée pour le corpus Turk-test en anglais, mais à une échelle moindre. Les scores du SARI sur le corpus self-test pour le français présentent une tendance opposée, c'est-à-dire que le score augmente avec le niveau de simplicité, à l'exception du score pour $0-1$. Contrairement à BLEU, SARI présente une variabilité moindre entre les niveaux de simplicité, mais ce score seul n'est pas assez concluant pour définir la simplicité du texte. Les scores FRE, à l'inverse, donnent une idée complète du niveau de simplicité.

Pour approfondir ces résultats, nous donnons sur la figure 2 la longueur moyenne des phrases pour chaque paire de niveaux. Nous observons des tendances claires : ainsi la longueur moyenne des phrases diminue à mesure que le niveau de simplification augmente. Cela s'observe pour presque tous les niveaux.

La tendance à la baisse est à nouveau évidente pour 2-$\{3, 4\}$. La longueur moyenne de phrase la plus faible concerne les simplifications produites par le système formé en utilisant un corpus de paires de niveaux 0-4, qui est la forme la plus simple.

Nous avons enfin réalisé des évaluations humaines des simplifications produites et avons observé que seule la sortie relative à des niveaux consécutifs de simplification est acceptable. On observe le même phénomène dans les scores BLEU. Les phrases dans les parties les plus simples, par exemple 1-3 et 1-4, sont vraiment très concises du côté simple et les modèles sont incapables d'apprendre les schémas de simplification complexes étant donné les intrications de suppression et de résumé impliquées dans le processus. Cependant, pour les niveaux consécutifs, par exemple 1-2, 2-3 et 3-4, la simplification était acceptable, ce qui indique qu'une simplification automatique basée sur un corpus synthétique est une démarche viable. Plus important encore, la qualité des systèmes appris avec de l'anglais est comparable à celle des systèmes appris avec du français synthétique et ces systèmes montrent un comportement similaire, ce qui tend à montrer que l'utilisation de corpus de simplification synthétique est une démarche viable pour les situations dans lesquelles les ressources sont peu nombreuses.

Références

ABDUL RAUF S., SCHWENK H., LAMBERT P. & NAWAZ M. (2016). Empirical use of information retrieval to build synthetic data for SMT domain adaptation. *ACM/IEEE Transactions on Audio, Speech and Language Processing (TASLP)*, **24**(4), 745–754.

ALVA-MANCHEGO F., BINGEL J., PAETZOLD G., SCARTON C. & SPECIA L. (2017). Learning how to simplify from explicit labeling of complex-simplified text pairs. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1 : Long Papers)*, p. 295–305.

APROSIO A. P., TONELLI S., TURCHI M., NEGRI M. & DI GANGI M. A. (2019). Neural text simplification in low-resource conditions using weak supervision. In *Proceedings of the Workshop on Methods for Optimizing and Evaluating Neural Language Generation*, p. 37–44.

BARBU E., MARTÍN-VALDIVIA M. T., MARTÍNEZ-CÁMARA E. & UREÑA-LÓPEZ L. A. (2015). Language technologies applied to document simplification for helping autistic people. *Expert Systems with Applications*, **42**(12), 5076–5086.

BROUWERS L., BERNHARD D., LIGOZAT A.-L. & FRANÇOIS T. (2014). Syntactic sentence simplification for french. In *Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR) @ EACL 2014*, p. 47–56, Gothenburg, Sweden.

BURLOT F. & YVON F. (2018). Using monolingual data in neural machine translation : a systematic study. In *Proceedings of the Third Conference on Machine Translation*, p. 144–155, Belgium, Brussels : Association for Computational Linguistics. DOI : 10.18653/v1/W18-64015.

CANNING Y. & TAIT J. (1999). Syntactic simplification of newspaper text for aphasic readers. In *ACM SIGIR'99 Workshop on Customised Information Delivery*, p. 6–11.

DAELEMANS W., HÖTHKER A. & SANG E. F. T. K. (2004). Automatic sentence simplification for subtitling in dutch and english. In *LREC*.

DE BELDER J. & MOENS M.-F. (2010). Text simplification for children. In *Proceedings of the SIGIR workshop on accessible search systems*, p. 19–26 : ACM ; New York.

EVANS R., ORĂSAN C. & DORNESCU I. (2014). An evaluation of syntactic simplification rules for people with autism. In *Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)*, p. 131–140, Gothenburg, Sweden : Association for Computational Linguistics. DOI : 10.3115/v1/W14-1215.

FAJARDO I., TAVARES G., ÁVILA V. & FERRER A. (2013). Towards text simplification for poor readers with intellectual disability : When do connectives enhance text cohesion ? *Research in developmental disabilities*, **34**(4), 1267–1279.

FLESCH R. (1948). A new readability yardstick. *Journal of applied psychology*, **32**(3), 221.

GRABAR N. & CARDON R. (2018). CLEAR – simple corpus for medical French. In *Proceedings of the 1st Workshop on Automatic Text Adaptation (ATA)*, p. 3–9, Tilburg, the Netherlands : Association for Computational Linguistics. DOI : 10.18653/v1/W18-7002.

HUENERFAUTH M., FENG L. & ELHADAD N. (2009). Comparing evaluation techniques for text readability software for adults with intellectual disabilities. In *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*, p. 3–10 : ACM.

K. PAPINENI S. R. & WARD R. (1998). Maximum likelihood and discriminative training of direct translation models. In *In Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, p. 189–192.

KANDEL L. & MOLES A. (1958). Application de l'indice de flesch à la langue française. *Cahiers Etudes de Radio-Télévision*, **19**(1958), 253–274.

KLEIN G., KIM Y., DENG Y., SENELLART J. & RUSH A. (2017). OpenNMT : Open-source toolkit for neural machine translation. In *Proceedings of ACL 2017, System Demonstrations*, p. 67–72, Vancouver, Canada : Association for Computational Linguistics.

KORHONEN A., TRAUM D. & MÀRQUEZ L. (2019). Proceedings of the 57th annual meeting of the association for computational linguistics. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*.

LAMBERT P., SCHWENK H., SERVAN C. & ABDUL-RAUF S. (2011). Investigations on translation model adaptation using monolingual data. In *Proceedings of the Sixth Workshop on Statistical Machine Translation*, p. 284–293, Edinburgh, Scotland : Association for Computational Linguistics.

MCCARTHY J. E. & SWIERENGA S. J. (2010). What we know about dyslexia and web accessibility : a research review. *Universal Access in the Information Society*, **9**(2), 147–152.

MOORE R. C. (2002). Fast and Accurate Sentence Alignment of Bilingual Corpora. In S. D. RICHARDSON, Éd., *Proc. Association for Machine Translation in America (AMTA O2)*, Lecture Notes in Computer Science 2499, p. 135–144, Tiburon, CA, USA : Springer Verlag.

NARAYAN S. & GARDENT C. (2014). Hybrid simplification using deep semantics and machine translation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, volume 1, p. 435–445.

NISIOI S., ŠTAJNER S., PONZETTO S. P. & DINU L. P. (2017). Exploring neural text simplification models. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)*, volume 2, p. 85–91.

PAPINENI K., ROUKOS S., WARD T. & ZHU W.-J. (2002). BLEU : a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the association for computational linguistics*, p. 311–318 : Association for Computational Linguistics.

RELLO L., BAEZA-YATES R., BOTT S. & SAGGION H. (2013a). Simplify or Help ? Text Simplification Strategies for People with Dyslexia. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*, W4A '13, New York, NY, USA : Association for Computing Machinery. DOI : 10.1145/2461121.2461126.

RELLO L., BAEZA-YATES R., DEMPERE-MARCO L. & SAGGION H. (2013b). Frequent words improve readability and short words improve understandability for people with dyslexia. In *IFIP Conference on Human-Computer Interaction*, p. 203–219 : Springer.

SCARTON C., PAETZOLD G. & SPECIA L. (2018). Text simplification from professionally produced corpora. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

SCARTON C. & SPECIA L. (2018). Learning simplifications for specific target audiences. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)*, p. 712–718.

SENNRICH R., HADDOW B. & BIRCH A. (2016). Edinburgh neural machine translation systems for WMT 16. In *Proceedings of the First Conference on Machine Translation : Volume 2, Shared Task Papers*, p. 371–376, Berlin, Germany : Association for Computational Linguistics. DOI : 10.18653/v1/W16-2323.

SURYA S., MISHRA A., LAHA A., JAIN P. & SANKARANARAYANAN K. (2019). Unsupervised neural text simplification. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, p. 2058–2068.

WUBBEN S., VAN DEN BOSCH A. & KRAHMER E. (2012). Sentence simplification by monolingual machine translation. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics : Long Papers-Volume 1*, p. 1015–1024 : Association for Computational Linguistics.

XU W., CALLISON-BURCH C. & NAPOLES C. (2015). Problems in current text simplification research : New data can help. *Transactions of the Association of Computational Linguistics*, **3**(1), 283–297.

ZHANG X. & LAPATA M. (2017). Sentence simplification with deep reinforcement learning. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, p. 584–594.

ZHAO S., MENG R., HE D., SAPTONO A. & PARMANTO B. (2018). Integrating transformer and paraphrase rules for sentence simplification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, p. 3164–3173.

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 10, p. 1353–1361, Beijing, China.