

Translating from Original to Simplified Sentences using Moses: When does it Actually Work?

Sanja Štajner

Research Group in Computational Linguistics
University of Wolverhampton, UK
SanjaStajner@wlv.ac.uk

Horacio Saggion

TALN Research Group
Universitat Pompeu Fabra, Spain
horacio.saggion@upf.edu

Abstract

In recent years, several studies have approached the Text Simplification (TS) task as a machine translation (MT) problem. They report promising results in learning how to translate from ‘original’ to ‘simplified’ language using the standard phrase-based translation model. However, our results indicate that this approach works well only when the training dataset consists mostly of those sentence pairs in which the simplified sentence is already very similar to its original. Our findings suggest that the standard phrase-based approach might not be appropriate to learn strong simplifications which are needed for certain target populations.

1 Introduction

Text Simplification (TS) aims to convert complex texts into simpler variants which are more accessible to a wider audiences, e.g. non-native speakers, children, and people diagnosed with intellectual disability, autism, aphasia, dyslexia or congenital deafness. In the last twenty years, many automatic text simplification systems have been proposed, varying from rule-based, e.g. (Brouwers et al., 2014; Saggion et al., 2015) to data-driven, e.g. (Zhu et al., 2010; Woodsend and Lapata, 2011), and hybrid (Siddharthan and Angrosh, 2014). Since 2010, there have been several attempts to approach TS as a machine translation (MT) problem (Specia, 2010; Coster and Kauchak, 2011a; Štajner, 2014). Instead of translating sentences from one language to another, the goal of text simplification is to translate sentences from ‘original’ to ‘simplified’ language.

In this paper, we seek to explore the main reasons for the success or failure of the phrase-based statistical machine translation (PB-SMT)

approach to TS. The results of our translation experiments in three languages indicate that the size of the dataset might not be the key factor for the success of this approach and that the effectiveness of such systems heavily depends on the similarity between the original and manually simplified sentences in the datasets used for training and tuning.

2 Related Work

Specia (2010) achieves BLEU score of 60.75 on a small (only 4,483 sentence pairs) dataset in Brazilian Portuguese, using the standard phrase-based translation model (Koehn et al., 2003) in the Moses toolkit (Koehn et al., 2007). The dataset consists of original sentences and their corresponding manually simplified versions obtained under the PorSimples project (Aluísio and Gasperin, 2010) following specific guidelines.

Coster and Kauchak (2011a) exploit the same translation model to learn how to simplify English sentences using 137,000 sentence pairs from Wikipedia and Simple English Wikipedia. They show that those results (BLEU = 59.87) can be improved by adding phrasal deletion to the probabilistic translation model, reaching the BLEU score of 60.46. Both those approaches seem to outperform all previous non-MT approaches to TS for English.

The fact that Specia (2010) and Coster and Kauchak (2011a) achieve similar performances of the PB-SMT system in spite of large differences in size of their datasets motivates our hypothesis that the key factor for a success of such an approach to TS might not lie in the size of the datasets but rather in the nature of the sentence pairs used for training and tuning of the PB-SMT models.

3 Methodology

We apply the following methodology:

- We run MT-based text simplification exper-

iments on three different datasets and languages following the methods proposed in previous studies (Specia, 2010; Coster and Kauchak, 2011a).

- We perform automatic evaluation in terms of the document-wise (BLEU) and the sentence-wise BLEU score (S-BLEU).
- We conduct a manual error analysis of the output of all three translation experiments.
- We calculate sentence-wise BLEU score on the training and development datasets to further understand the differences observed in the translation experiments.

3.1 Datasets

We use three sentence-aligned TS corpora in three different languages:

1. **EsSim** – The corpus of original news texts in Spanish and their manual simplifications aimed at people with Down syndrome. Simplification was performed by trained human editors under the Simplext project (Saggion et al., 2015).
2. **PorSim** – The corpus of original news texts in Brazilian Portuguese and their manual simplifications compiled under the PorSimples project (Caseli et al., 2009). Original sentences and their corresponding ‘natural’ simplifications of this corpus were used for in the previous PB-SMT experiments (Specia, 2010).
3. **Wiki** – The parallel corpus of automatically aligned sentence pairs from English Wikipedia and Simple English Wikipedia, used for the PB-SMT experiments by Coster and Kauchak (2011a).¹

In order to compare the results of translation experiments among the three corpora, we train and tune all three systems on a similar amount of data. Therefore, we focus only a subset of sentence pairs used by Specia (2010), and by Coster and Kauchak (Coster and Kauchak, 2011a). The sizes of the corpora used are shown in Table 1.

¹<http://www.cs.middlebury.edu/~dkauchak/simplification/>

	EsSim	Wiki	PorSim
Training	745	800	800
Dev.	90	200	200
Test	90	100	100
Total	925	1100	1100
Selection	all	random	random

Table 1: Size of the corpora

3.2 Translation Experiments

We run three MT experiments using the standard PB-SMT models (Koehn et al., 2003) implemented in the Moses toolkit (Koehn et al., 2007) and GIZA++ (Och and Ney, 2003) to obtain the word alignment. The English experiment uses the Wiki aligned corpus for translation model (TM) and the English part of the Europarl corpora² for building the language model (LM). The Spanish experiment uses the EsSim dataset to build the TM and the Spanish Europarl for the LM. The Brazilian Portuguese experiment uses the PorSim dataset for the TM and the Lácio-Web corpus³ in Brazilian Portuguese for the LM⁴. The sentence pairs for training, development and test sets are selected randomly from the initial dataset.

4 Results and Discussion

In the next three subsections, we present and discuss the results of the automatic evaluation of the translation experiments (Section 4.1), the error analysis of the translation experiments (Section 4.2), and the distribution of the S-BLEU score across the four datasets (Section 4.3).

4.1 Automatic Evaluation

The results of the translation experiments and sentence similarity metrics on the three datasets used for training the translation models are presented in Table 2. The BLEU scores achieved by translation experiments in English and Brazilian Portuguese are similar to those reported by Specia (2010) and Coster and Kauchak (2011a) in spite of our experiments having reduced the sizes of the two corpora for fair comparison with the Spanish dataset. As can be observed (Table 2), we cannot claim to

²<http://www.statmt.org/europarl/>

³<http://www.nilc.icmc.usp.br/lacioweb/>

⁴The Portuguese in the Europarl corpora belong to the different regional language variety, and thus we opted for the Lácio-Web corpus written in the same regional variety as the used TS dataset.

	EsSim	Wiki	PorSim
BLEU	10.55	53.28	65.66
t-BLEU	10.16	56.39	48.46
BP	0.59	0.87	0.93
S-BLEU	0.16	0.58	0.58

Table 2: Automatic evaluation

have an equally good performance on the Spanish dataset for which we obtained a BLEU score of 10.55.

In order to understand better the differences in translation performances (BLEU) across datasets, we calculated BLEU score (t-BLEU) with brevity penalty (BP), and sentence-wise BLEU score (S-BLEU)⁵ on the training datasets (EsSim, PorSim, and Wiki). The manual simplifications (or in the case of English, the Simple Wikipedia versions) were used as hypotheses and the original non-simplified versions as references. It appears that the similarity between the original and simplified sentences used for training is much higher (up to four times higher in the case of the *S-BLEU*) in the Wiki and PorSim datasets than in the third dataset (EsSim).

It can be noted that the EsSim dataset achieves significantly lower BLEU score than the other two. Additionally, the EsSim dataset has a much higher brevity penalty (BP) on the training set than the other two datasets, indicating that the sentence shortening is more commonly used simplifying operation in this dataset than in the other two. It seems that whenever MT performs well (Table 2), we actually have a dataset that is more MT-looking and complies with the underlying assumptions of the standard phrase-based model (reflected in the high BLEU score on the training data). The low BLEU score on the training dataset (t-BLEU) suggests that there are many string transformations and strong paraphrases to be learnt, and thus the standard phrase-based translation model might not be the most suitable for the task.

As it is known that the BLEU score does not give a fair comparison among systems with different architectures – or, in this case, systems trained for different languages and tested on different datasets – we do not rely on the automatic

⁵Sentence-level BLEU score (S-BLEU) differs from BLEU score only in the sense that S-BLEU will still positively score segments that do not have higher n-gram matching (n=4 in our setting) unless there is no unigram match; otherwise it is the same as BLEU.

Modification	EsSim	PorSim	Wiki
None	4.44%	40%	65%
1 Substitution	4.44%	40%	28%
>1 Substitution	91.11%	20%	2%
Split	6.67%	14%	5%
Combined	6.67%	14%	3%

Table 3: Classification of modifications

evaluation of our models. Instead, we perform a detailed manual analysis of the output of all three systems.

4.2 Error Analysis

In order to clarify doubts raised by the results of the automatic evaluation, we performed error analysis on all sentences from the three datasets (90 sentences in Spanish, 100 sentences in English, and 100 sentences in Portuguese). The classification of the test sentences based on the number and type of modifications made by the translation/simplification models is presented in Table 3.

The manual examination of the output of the translation model trained on the EsSim corpora confirmed the poor performance of the system, describing the output of the automatic simplification as barely comprehensible in most of the cases. At the same time, the manual examination of the output of the model trained on the Wiki corpus revealed that such a high performance (in terms of the BLEU score) is indeed due to the fact that the output was in most cases identical to the original sentence. The human evaluation of the automatically generated sentences in Portuguese, confirmed that the model was able to perform at least one lexical substitution in 40% of the sentences, more than one lexical substitution in 20% of the cases, and sentence splitting combined with lexical substitution(s) in 14% of the cases.

It was not possible to perform usual human evaluation of the generated sentences in terms of their grammaticality, simplicity, and meaning preservation due to: (1) 82 automatically simplified sentences in Spanish (out of 90) being severely ungrammatical (due to the multiple wrong word substitutions), to the extent that it was impossible to assess their simplicity (as they were barely comprehensible) or to compare their meaning with the meaning of the original sentence, and (2) 65 automatically simplified sentences for English being identical to their originals and another

Ex.	Original	Automatic	Reference
p1	Porém , ninguém pode subestimar a capacidade de mobilização de Chávez.	Mas ninguém pode subestimar a capacidade de mobilização de Chávez.	Mas ninguém pode <i>desprezar</i> a capacidade de mobilização de Chávez.
p2	Dom Odilo, que trabalhou no Vaticano entre 1994 e 2001, é tido como um religioso moderado, próximo à Cúria Romana e alinhado com as diretrizes da Santa Sé.	Dom Odilo trabalhou no Vaticano entre 1994 e 2001. Ele é tido como um religioso moderado, próximo à Cúria Romana e concorda com as normas da Santa Sé.	Dom Odilo trabalhou no Vaticano entre 1994 e 2001. Ele é tido como um religioso moderado, próximo à Cúria Romana e <i>que</i> concorda com as normas da Santa Sé.
p3	A segunda sugestão de modificação diz respeito ao veto de pesquisas com células-tronco embrionárias humanas e com clonagem terapêutica .	A segunda sugestão de modificação diz respeito à proibição de pesquisas com células-tronco de embriões humanos e com clonagem para tratamento de doentes .	A segunda sugestão de modificação diz respeito à proibição de pesquisas com células-tronco de embriões humanos e com clonagem para tratamento de doentes.

Table 4: Examples of the automatic simplification in Brazilian Portuguese (differences between the original sentences and their automatic simplifications are shown in bold, and the deviations of the manual simplifications from the automatic simplifications are shown in italics)

28 sentences differing from their originals by only one word. Therefore, we focused on detailed analysis of the generated sentences in all three languages, seeking to discover what are the possibilities and limitations of our simplification models.

4.2.1 Portuguese

Table 4 shows examples of the original sentences from the test dataset (*Original*), their automatic simplifications (*Automatic*), and their corresponding reference simplifications (‘gold standards’) manually simplified under the PorSimples project (*Reference*). As previously mentioned, 60 out of 100 original test sentences were lexically modified by the system, while 14 of them were additionally split into two sentences.

In the first example (*p1*), the system performed one lexical substitution replacing the word “Porém” (*however*) with “Mas” (*but*). The same substitution was done by human editors. However, the system only performed this one substitution, while the manual simplification encompassed one additional lexical simplification.

In the second example (*p2*), the system performed a correct sentence splitting taking the apposition in a separate sentence (“Dom Odilo trabalhou no Vaticano entre 1994 e 2001.”), and two correct lexical simplifications: “alinhado” (*aligned, in line*) was changed into “concorda” (*agree, comply*) and “diretrizes” (*guidelines*) into “normas” (*standards, norms*). The difference between the manual and automatic simplification of this sentence was not significant (the automatically simplified sentence is still grammatical, although the manually simplified sentence might be

stylistically better).

The third example (*p3*) shows a case in which the automatic simplification managed to reach the level of manual simplification by performing three corrected lexical simplifications and generating the output sentence equal to the manually simplified sentence.

4.2.2 Spanish

Table 5 shows examples of the original sentences from the test dataset (*Original*), their automatic simplifications (*Automatic*), and their corresponding reference simplifications (‘gold standards’) manually simplified under the Simplext project (*Reference*).

In the first example (*s1*), “UE” (*EU*) was correctly replaced with “Europa” (*Europe*), while the incorrect substitution of “sacar de la pobreza” (*get out of poverty*) with “objetivo” (*goal/aim/objective*) left the sentence meaningless. Together with the deletion of “20” (in “20 million people”) and “hasta 2020” (*until 2020*), and the insertion of “a” at the end of the sentence, the generated sentence is completely ungrammatical and meaningless. The original sentence “*The EU wants to get out of poverty 20 million people until 2020*” is simplified as “*The Europe wants goal to millions of people*”.

The second example (*s2*) is particularly interesting as the manual simplification (‘gold standard’) is identical to the original sentence. In the automatically generated sentence, however, the phrase “dimitirá como presidente” (*will quit as a president*) in the original sentence was correctly ‘translated’ as “deja la presidencia” (*leaves the presi-*

Ex.	Original	Automatic	Reference
s1	La UE quiere sacar de la pobreza a 20 millones de personas hasta 2020 .	La Europa quiere objetivo a millones de personas a .	Europa quiere ayudar a millones de personas a dejar de ser pobres.
s2	Alex de la Iglesia dimitirá como presidente de la Academia de Cine.	Cine Alex de la Iglesia deja la presidencia de la Academia de Cine.	Alex de la Iglesia dimitirá como presidente de la Academia de Cine.
s3	Por otro lado , el informe de “la Caixa” sitúa en el 9,5% del PIB el déficit público al término de 2010 y calcula que el conjunto de las administraciones podrán reducirlo hasta el 6,4% durante 2011.	, el informe de la sitúa en el 9,5 mitad el el déficit público al término de 2010 y creo que el literaria de las corridas podrán reducirlo hasta el 6,4 mitad en 2011.	El informe también mostraba que una parte de la deuda en 2010 era de las administraciones públicas. El informe calcula que las administraciones públicas podrán reducir su deuda en 2011.

Table 5: Examples of the automatic simplification in Spanish (differences between the original sentences and the automatically simplified sentences are given in bold)

Ex.	Original	Automatic	Reference
e1	Parennes is a commune in the Sarthe department in the region of Pays-de-la-Loire in north-western France.	Parennes is a commune. It is found in the Sarthe department in the region Pays de la Loire in north-western France.	Parennes is a commune. It is found in the <i>region Pays de la Loire in the Sarthe department in the west of France</i> .
e2	Mated females will lay a clutch of 30-60 eggs in a mound of dry leaves or peat.	Mated women will lay a clutch of 30-60 eggs in a mound of dry leaves or peat.	Mated <i>females</i> will lay a clutch of 30-60 eggs in a mound of dry leaves or peat.
e3	Elections to the new authorities were held on three different Thursdays in 1973.	Elections to the new government were held on three different Thursdays in 1973.	<i>Elections were held to the new authorities</i> on three different Thursdays in 1973.

Table 6: Examples of the automatic simplification in English (differences between the original sentences and their automatic simplifications are shown in bold, and the deviations of the manual simplifications from the automatic simplifications are shown in italics)

dency). One could argue that the phrase used in the automatically simplified sentence is actually simpler than the corresponding phrase in the ‘gold standard’ (and the original), as the verb “dejar” (*to leave*) is more frequent than the verb “dimitir” (*to quit*). This complies with the common practice in text simplification to replace the infrequent and more specific terms/phrases with their more frequent synonyms. The native speakers might argue that use of the verb “dejar” (*to leave*) introduces ambiguity (as it is not clear whether Alex leaves his presidency because his mandate is over or because he is quitting), while the use of the verb “dimitir” (*quit*) does not leave any doubt about the way/reason Alex is leaving his presidency. Still, non-native speakers will definitely be familiar with the Spanish word “dejar”, while (depending on their level of Spanish) may not be familiar with the Spanish word “dimitir”.

The third example (s3) represents one of the most frequently observed cases of automatic simplification in the test dataset. In those cases, the PB-SMT system generates the output which is at the same time ungrammatical (mostly due to the incorrect deletions of various sentence parts) and meaningless (mostly due to the incorrect word

substitutions, but also due to the ungrammatical sentence constructions). For instance, the word “conjunto” (*set*) is replaced with the word “literaria” (*literary*), and the word “administraciones” (*administrations*) with the word “corridas” (*runs*). In the first case, the original word was replaced with the word with a different part-of-speech (a noun replaced with an adjective). However, this example (s3) also shows a particularly interesting case of lexical simplification performed by the PB-SMT system, but not performed by the human editor. The word “calcula” (*calculates*) is replaced with the word “creo” (*believes*). In this sentence, the word “calcula” (*calculates*) was indeed used with the meaning “creo” (*believes*), which is not its most common meaning. Such replacements are favourable in text simplification, as stated in Web Content Accessibility Guidelines (W3C, 2008).

4.2.3 English

Table 6 contains several examples of the original sentences from the test dataset (*Original*), their automatic simplifications (*Automatic*), and their corresponding reference simplifications (‘gold standards’) from the Simple English Wikipedia (*Reference*). They illustrate some of the phenomena

Corpus	[0, 0.3)	[0.3, 0.4)	[0.4, 0.5)	[0.5, 0.6)	[0.6, 0.7)	[0.7, 0.8)	[0.8, 0.9)	[0.9, 1]
EsSim	85.96%	4.45%	1.62%	0.94%	0.94%	0.27%	0.40%	5.40%
PorSim	12.96%	11.20%	11.74%	18.08%	13.23%	12.82%	7.83%	12.28%
Wiki	26.86%	6.48%	9.31%	6.34%	8.37%	6.88%	6.75%	29.15%

Table 7: Distribution of the S-BLEU scores (columns represent the intervals for S-BLEU)

revealed during the manual error analysis.

Example *e1* presents one of the five correctly performed sentence splittings learned by the PB-SMT system. However, it is important to mention that all five split sentences in the test dataset share the same structure of the original sentence (*'X is a commune in...'*). In all five cases, such an original sentence is transformed into two sentences which again share the same structure (*'X is a commune. It is found in...'*). The example *e2* presents an example of a bad word substitution (lexical simplification which leads to a simpler sentence but changes the original meaning), while *e3* shows a good word substitution (lexical simplification).

It can be noted that all examples of the automatically simplified sentences are still grammatical. One or two wrongly applied word substitutions may only change the meaning of the sentence but they do not deteriorate the grammaticality of the sentence. Correctly applied word substitutions and sentence splittings preserve the original meaning and grammaticality of the sentence, and lead to a slightly simpler output.

4.3 Distribution of S-BLEU Scores

A closer examination of the S-BLEU distribution (Table 7) indicate that the cause behind the good performance of the ‘translation’ system trained on PorSim and Wiki datasets probably lies in the nature of the data. The Wiki corpus contains only those sentence pairs whose normalised similarity was higher than 0.5 (Coster and Kauchak, 2011b). The PorSim corpus consists only of the sentence pairs simplified by ‘natural’ simplification in which the most common simplifying operation is sentence splitting (Gasperin et al., 2009). EsSim corpus, on the other hand, contain a great number of deletions and strong paraphrases (combinations of lexical and syntactic transformations with deletions) as reported by Štajner *et al.* (2013). Such strong paraphrases and reordering of clauses within a sentence are very frequent in the EsSim dataset, while hardly present in the Wiki and Por-

Sim datasets.⁶ Although well-motivated and necessary for the target population in mind (people with intellectual impairments), those transformations cannot be learnt by the standard PB-SMT model.

5 Conclusions and Future Work

Text simplification has recently been treated as a statistical machine translation problem. By comparing the performance of this translation paradigm across three datasets, we have identified possible causes for the success and failure of such a simplification approach. It appears that learning how to ‘translate’ from original to simplified language using standard PB-SMT model works well only in some special cases, when the training data mostly consists of the sentence pairs which are already very similar.⁷ Our results indicate that this approach would not be effective if we want to learn ‘real’, strong simplifications like those performed by trained human editors familiar with the specific needs of their target population (e.g. people with intellectual disabilities). Those simplifications involve linguistically rich transformations (e.g. paraphrase, summarisation) which cannot be modelled by standard PB-SMT systems.

We are currently investigating how to improve the translation model with the addition of synonym datasets and the language model using a large bootstrapped corpus of “simple” sentences instead of normal, non-simplified language.

Acknowledgements

The research described in this paper was partially funded by the project SKATER-UPF-TALN (TIN2012-38584-C06-03), Ministerio de Economía y Competitividad, Secretaría de Estado de Investigación, Desarrollo e Innovación, Spain,

⁶For examples from all three corpora and a more detailed discussion see (Štajner, 2015).

⁷Our recent study on PB-SMT approach to text simplification using larger datasets for English (Štajner et al., 2015) confirms these findings.

and the project ABLE-TO-INCLUDE (CIP-ICT-PSP-2013-7/621055). We would also like to thank Hannah Béchara, Wilker Aziz, and Lucia Specia for the help with the translation experiments and resources.

References

- Sandra Maria Aluísio and Caroline Gasperin. 2010. Fostering digital inclusion and accessibility: The porsimples project for simplification of portuguese texts. In *Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas, YIW-CALA '10*, pages 46–53, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Laetitia Brouwers, Delphine Bernhard, Anne-Laure Ligozat, and Thomas François. 2014. Syntactic sentence simplification for french. In *Proceedings of the EACL Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)*, Gothenburg, Sweden, pages 47–56.
- Helena M. Caseli, Tiago F. Pereira, Lucia Specia, Thiago A. S. Pardo, Caroline Gasperin, and Sandra M. Aluísio. 2009. Building a Brazilian Portuguese parallel corpus of original and simplified texts. In *10th Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2009)*.
- William Coster and David Kauchak. 2011a. Learning to Simplify Sentences Using Wikipedia. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 1–9.
- William Coster and David Kauchak. 2011b. Simple English Wikipedia: a new text simplification task. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, Oregon, USA: Association for Computational Linguistics*, pages 665–669.
- Caroline Gasperin, Lucia Specia, Tiago F. Pereira, and Sandra M. Aluísio. 2009. Learning When to Simplify Sentences for Natural Text Simplification. In *Proceedings of the Encontro Nacional de Inteligncia Artificial (ENIA-2009)*, Bento Gonalves, Brazil., pages 809–818.
- Philipp Koehn, Franz Josef Och, and Daniel Marcu. 2003. Statistical phrase-based translation. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1, NAACL '03*, pages 48–54, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin, and Evan Herbst. 2007. Moses: Open source toolkit for statistical machine translation. In *Proceedings of ACL*.
- Franz Josef Och and Hermann Ney. 2003. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1):19–51.
- Horacio Saggion, Sanja Štajner, Stefan Bott, Simon Mille, Luz Rello, and Biljana Drndarevic. 2015. Making It Simplext: Implementation and Evaluation of a Text SImplification System for Spanish. *ACM Transactions on Accessible Computing (TACCESS)*, 6(4).
- Advaith Siddharthan and M.A. Angrosh. 2014. Hybrid text simplification using synchronous dependency grammars with hand-written and automatically harvested rules. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Gothenburg, Sweden, pages 722–731.
- Lucia Specia. 2010. Translating from complex to simplified sentences. In *Proceedings of the 9th international conference on Computational Processing of the Portuguese Language*, pages 30–39, Berlin, Heidelberg.
- Sanja Štajner, Biljana Drndarević, and Horacio Saggion. 2013. Corpus-based Sentence Deletion and Split Decisions for Spanish Text Simplification. *Computación y Sistemas*, 17(2):251–262.
- Sanja Štajner, Hannah Bechara, and Horacio Saggion. 2015. A Deeper Exploration of the Standard PB-SMT Approach to Text Simplification and its Evaluation. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 823–828, Beijing, China. ACL.
- Sanja Štajner. 2014. Translating sentences from ‘original’ to ‘simplified’ spanish. *Procesamiento del Lenguaje Natural*, 53:61–68.
- Sanja Štajner. 2015. *New Data-Driven Approaches to Text Simplification*. Ph.D. thesis, University of Wolverhampton, UK.
- W3C, 2008. *Web Content Accessibility Guidelines (WCAG) 2.0*.
- Kristian Woodsend and Mirella Lapata. 2011. Learning to Simplify Sentences with Quasi-Synchronous Grammar and Integer Programming. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Zhemín Zhu, Delphine Berdard, and Iryna Gurevych. 2010. A Monolingual Tree-based Translation Model for Sentence Simplification. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 1353–1361.