Translating from Complex to Simplified Sentences

Lucia Specia

Research Group in Computation Linguistics
University of Wolverhampton
Wolverhmapton, UK
L.Specia@wlv.ac.uk

Abstract. We address the problem of simplifying Portuguese texts at the sentence level by treating it as a "translation task". We use the Statistical Machine Translation (SMT) framework to learn how to translate from complex to simplified sentences. Given a parallel corpus of original and simplified texts, aligned at the sentence level, we train a standard SMT system and evaluate the "translations" produced using both standard SMT metrics like BLEU and manual inspection. Results are promising according to both evaluations, showing that while the model is usually overcautious in producing simplifications, the overall quality of the sentences is not degraded and certain types of simplification operations, mainly lexical, are appropriately captured.

Keywords: Text Simplification, Statistical Machine Translation, Portuguese

1 Introduction and background

The ability to use written language to obtain and record information, express themselves, plan and learn continuously is called *literacy*, or "letramento", in Brazil. According to the index used to measure the literacy level of the population in Brazil (*INAF - National Indicator of Functional Literacy*), a vast number of people belong to the so called *rudimentary* and *basic* literacy levels [1]. These people are able to find explicit information in short texts (rudimentary level) and also process slightly longer texts and make simple inferences (basic level).

The PorSimples project[1] aims at producing tools to simplify complex texts and therefore promote digital inclusion and accessibility for people with such levels of literacy, and possibly other kinds of reading disabilities. Amongst the tools envisaged is a browser plugin that automatically simplifies texts on the web for the end-user. The focus is on texts published in government sites or by relevant news agencies, both expected to be of importance to a large audience with various literacy levels. The language of the texts is Brazilian Portuguese, for which there are no text simplification systems, to the best of our knowledge.

Text Simplification (TS) has been exploited in other languages for helping poor literacy readers [2], [3] and [4] and people with disabilities such as aphasia [5]. It has

[1] http://caravelas.icmc.usp.br/wiki/index.php/English

also been used for improving the accuracy of other natural language processing tasks [6, 7]. Most approaches focus on developing rule-based systems based on parser structures and limited to certain simplification operations.

Corpus-based systems, which can learn the simplification operations, have also been exploited through the use of parallel aligned corpora of original and simplified texts. [8] investigates the induction of syntactic rules from a corpus by extracting syntactic correspondences and generalizing them. [9] uses corpora of TV program transcripts and subtitles to generate subtitles for hearing-impaired people, focusing on summarization and lexical substitution. [10] uses machine learning techniques to aid second language teachers by proposing when to split or drop sentences.

In the scope of the PorSimples project, one of the main text processing strategies to produce simplified texts is to identify and apply lexical and syntactic simplification operations. A rule-based system based on a pre-defined set of operations has been implemented [11] and a machine learning algorithm has been used to support the decision of whether or not to simplify a sentence [12]. Both systems rely on the set of pre-defined simplification operations. These operations, described in Section 2, were carefully designed based on a study of the Portuguese grammar. Their general definition cover all phenomena believed to cause sentences to be difficult to read. However, the specific implementation of such operations as rules is not a trivial task. The rules are defined in terms of syntactic representations and lexical clues, and it is impossible to foresee all variations of the same phenomena with different realizations. Generalizing the rules while making sure they do not cover illicit phenomena is therefore a complex issue. Moreover, the dependency on a syntactic parser, which is not error-free, might incur inadequate uses of the rules. Finally, a very important limitation is the fact that this system can only cover the phenomena that fall into the pre-defined set of operations.

All the existing corpus-based approaches aim at the simplification of English texts, mostly to learn when to apply operations or to induce the actual simplification rules based on syntactic representations. In this paper we proposed an alternative strategy to exploit parallel text with no additional annotation. We learn how to translate from complex to simplified Portuguese texts using the Statistical Machine Translation (SMT) framework, or more specifically, the noisy-channel model as implemented in a standard phrase-based SMT system. The SMT framework has been used for other tasks than multilingual translation, including recasing (http://www.statmt.org/moses/) and other post-edition for problems [13], but not for Text Simplification.

In what follows, we first describe the experimental setting, including the corpus of original and simplified texts and the SMT framework (Section 2) to then present our experiments and results (Section 3).

2 Experimental setting

2.1 The parallel corpus of original and simplified texts

Once specified the set of syntactic and lexical simplification operations to be addressed by the PorSimples project, one of the main tasks was the manual creation,

by an expert in TS, native speaker of Portuguese, of a parallel corpus of original and simplified texts [14]. The pre-defined specifications aim to simplify texts as much as possible, targeting mainly the rudimentary level readers. However, this sometimes leads to oversimplification of the text, making it unnatural, tiresome or even disturbing for readers with basic (or higher) literacy level. Therefore, the Project also considers the case of performing simplification only if the resulting text still looks "natural". These two levels of simplification are denominated "strong" and "natural", respectively, and were taken into account during the creation of the parallel corpus. In this paper, we use the corpus of natural simplifications to train the "translation" system. This type of simplification is more difficult to be captured by rules, as it does not follow rigid constraints like those imposed for strong simplifications.

The pre-defined simplification operations are the following: (1) non-simplification; (2) lexical substitution to replace unusual or complex words; (3) simple rewriting (to replace discourse markers or sets of words, like idioms or collocations) (4) putting the sentence in its subject-verb-object order; (5) putting the sentence in the active voice; (6) inverting the clause ordering; (7) splitting or (8) joining sentences; (9) dropping the sentence or (10) dropping parts of the sentence. Besides the 10 regular operations, during the annotation process, the annotator could choose (11): strong rewriting, for any sort of free rewriting of sentence that was not covered by the other operations, as defined in [10].

The corpus was extracted from two of the main Brazilian newspapers: first page news articles from *Zero Hora* and science articles from *Folha de Sao Paulo*. The human annotator used a tool to perform the simplifications, which recorded, among other information, the alignment between original and simplified sentences and the type of operation used [14]. Table 1 shows examples of original sentences in (A) and (C), and their natural simplifications in (B) and (D). The operations applied to (A) to make (B) were regular ones (splitting the sentence and replacing unusual words). (D), on the other hand, was produced from (C) using "strong rewriting" for natural modifications not covered by the pre-defined set.

Table 1. An example of an original text (A) and its (natural) simplified version (B)

A	*A cerração, que ocultou o amanhecer dos porto-alegrenses, voltou a comprometer as operações do aeroporto por mais de oito horas, entre 1h50 min e 9h56 min.*
B	*A neblina comprometeu outra vez as operações do aeroporto por mais de oito horas. A neblina comprometeu as operações do aeroporto entre 1h50 min e 9h56 min. A neblina ocultou o amanhecer dos porto-alegrenses.*
C	*Cientistas britânicos detectaram, em adultos, a produção de células hepáticas a partir de células-tronco da medula óssea.*
D	*Cientistas britânicos detectaram em adultos que células-tronco da medula óssea produziram células do fígado.*

The resulting parallel corpus is composed of 104 news articles from *Zero Hora* and 37 from *Caderno da Ciencia*, amounting to 4,483 original sentences and their corresponding simplifications. Multiple sentences produced from a single complex sentence were put together and aligned to the original sentence. We call them *segments*. The distribution of the operations used in the manual simplification process

of the original sentences is depicted in Table 2. A detailed corpus study is presented in [14].

Table 2. Statistics on the types of simplification in the parallel corpus

Syntactic and Lexical Operations	Number of (original) sentences
Non-simplification	595
Lexical Substitution	1487
Simple rewriting	732
Subject-verb-object ordering	41
Transformation to active voice	100
Inversion of clause ordering	256
Splitting sentences	823
Joining sentences	12
Dropping one sentence	0
Dropping sentence parts	424
Strong rewriting	13

2.2 SMT framework and Moses

The standard framework for Statistical Machine Translation (SMT) is based on the noisy-channel model from information theory. The translation of a string *f* in a foreign language into a string *e* in a native language is based on estimations of the probability *p(e|f)*. The Bayes Theorem is applied to decompose this problem into two others:

$$p(e \mid f) = \frac{p(f \mid e)\, p(e)}{p(f)} \qquad (1)$$

where *p(f|e)*, called the *translation model*, is the probability that the foreign string is the translation of the native string, and *p(e)*, called the *language model* is the probability of seeing that native string. *p(f)* is a normalization constant and can be disregarded. Therefore finding the best translation ê becomes the problem of finding the translation with the highest probability:

$$\hat{e} = \arg\max_{e \in e^*} p(e \mid f) = \arg\max_{e \in e^*} p(f \mid e)\, p(e) \qquad (2)$$

The basic assumption of this paper is that if SMT works well for translating between two different languages, it could also succeed in translating from complex to simplified texts in the same language. The goal is to learn a probabilistic dictionary of phrases and their corresponding simplified versions (*translation model*) along with a model that indicates how likely it is that the segment is a good simplified segment in Portuguese (*language model*). Besides, it is necessary to estimate weights for these and additional models governing, for example, how to distribute such phrases in a simplified segment (*reordering model*), the penalty for source and translation with different lengths (*word penalty*), etc. Each candidate translation *t* for a given source

segment is scored according to a linear combination of all these models (*feature functions* f_j) and their weights λ_j learned during a tuning phase:

$$score(t) = \sum_j \lambda_j f_j(t) \qquad (3)$$

We use Moses, a standard SMT system [15] considered by many as the state of the art in phrase-based SMT. Moses is a generalization of the noisy channel, where the following feature functions are considered to cover the language and translation models[2]:

- phrase translation model: 5 features covering conditional forward and inverse phrase and lexical translation probabilities, and phrase penalty;
- target language model;
- distortion limit: a limit on the amount of reordering allowed;
- word penalty: to penalize translations that differ in length as compared to the source segment;
- lexicalized reordering model: 6 features to cover reordering of lexical items.

The weights of these features are estimated using *Minimum Error Rate Training* [16] based on some tuning data (source and reference translation). We use configuration of the system suggested in the series of shared translation tasks where Moses is used as the *baseline system*[3]. This is the best setting for most language-pairs and domains.

3 Experiments and results

For the set of experiments presented here we put together all the segments in Table 2. With that, we are targeting a very robust scenario in which given a certain sentence and no further information about it, the model should decide whether it needs to be simplified and, if so, perform the necessary simplifications. In that scenario, sentences that do not need simplification should simply be copied to the output.

From our corpus of 4,483 segment pairs, we randomly selected:

- 3,383 segments for *training* (i.e., for generating the dictionary of original-simplified phrases);
- 500 segments for *tuning* the parameters in Moses; and
- 500 segments for *test* (evaluation).

3.1 Training

The phrase-table produced from training contains 229,000 phrases. The default configuration in Moses builds phrases with up to 7 tokens. While many of them

[2] http://www.statmt.org/moses/?n=Moses.FeatureFunctions
[3] http://www.statmt.org/wmt09/baseline.html

simply have the same source and target phrases with high probability scores, others cover adequately many types of lexical simplification and simple rewriting. Some interesting examples are shown in Table 3, together with their main phrase probability feature *p(f|e)*, which is given the highest weight in the tuning phase (Section 3.2).

Table 3. Examples of phrased in our phrase table and their probabilities

Source phrase	Target phrase	*p(f|e)*
para	com o objetivo de	0.8
para	com o objetivo	0.709677
para a melhor compreensão da	com o objetivo de compreender melhor a	1
, quando	, momento em que	0.625
, quando	, no momento em que	0.5
, quando	, no momento	1
-- não gostamos quando	-- não gostamos se	1
segundo	de acordo com	0.266667
dias , segundo	dias de acordo com	1
“ (editora maranta , são paulo ,	rambelli	0.004
negocio ainda precisa ser	mas	0.0023
de acordo com	a imprensa reproduzisse	0.5

It is important to notice that many spurious phrases are also produced, such as the three last cases in Table 3. However, they either have low probabilities and are therefore ruled-out from the candidate translation set by the translation model, or are ruled-out by the other models (mainly word-penalty and language model).

3.2 Tuning

By inspecting the weights estimated for each feature from the tuning data, we found that the features with highest individual weight were word penalty (negative weight) and language model (positive weight). Absolute values for the weights of phrase probability and some lexicalized reordering features are also significantly higher than the remaining features. This is an indication that our model prioritizes simplifications which are not very different in length and order as compared to the original segment, and which are also common segments in the target (simplified) language.

3.3 Test

After building the phrase-table and tuning Moses, we decoded the 500 test segments and computed the two most common MT evaluation metrics: BLEU [17] (1 to 4-grams) and NIST [18] (1 to 5-grams). Both BLEU and NIST check the overlapping of n-grams between two sets of segments, in this case, human and machine simplification. There are no standard automatic metrics like these for Text Simplification.

Let *source* be a given original sentence, *target* the system simplification and *reference* the human simplification for that source sentence. The scores obtained for the target against the reference are **BLEU = 0.6075** and **NIST = 9.6244**. The upper

bound for BLEU and NIST with this corpus would be **1** and **~13**, respectively. Although it is impossible to compare such scores with those reported in standard translation tasks, a BLEU of ~0.60 in a translation task usually means good translation quality, particularly for News texts (the shared evaluation task reports scores of around 0.25 for translation from English into several languages) [19].

Using these evaluation metrics and also checking string matching (ignoring case, spaces and punctuations), we performed three other tests:

1) **Amount of simplification performed**: we check whether the target segments differ significantly from the source sentences. The higher the BLEU/NIST scores between these two versions, the closer the target segments are to the source. The scores obtained are **BLEU = 0.7664** and **NIST = 11.509**. This shows that the simplifications produced by the system are closer to the source than to the reference segments, which is an indication that the system is overcautious in performing simplifications. In other words, in terms of standard metrics recall and precision, the system favors precision. By measuring string matching, we found that **374 (75%)** segments in the target differ from the source, the remaining were not modified.
2) **Contribution of the simplifications performed**: we check whether the reference segments match better the target or source segments. The scores comparing source and reference segments are **BLEU = 0.5726** and **NIST = 9.5101.** This shows that target segments are closer to the reference (**BLEU = 0.6075** and **NIST = 9.6244**), which proves that the simplifications being performed are likely to be correct, since they approximate the segments to the gold-standard ones. However, exact matching is rare: **452 (90%)** segments in the target differ from the reference.
3) **Preservation of originally simple sentences**: we compare the string matching between source and reference segments, finding that that **443 (89%)** source sentences were modified somehow by the human annotator, and should, in principle, have also been simplified by our model. Out of these 443 segments, the SMT model only modified **350 (79%)**, which shows again that the system was overcautious. If we recall that in total the model modified 374 segments (item (1)), we see that only **24 segments** that were not simplified by the annotator were modified by our model.

Although this inspection gives some intuition on the performance of the model, automatic metrics like BLEU/NIST and string matching are known to be very limited even for standard translation evaluation tasks. We therefore also performed manual inspection of a subset of segments. We randomly selected 20 segments out of the 500-test set which suffered some modification and checked whether the simplified versions produced were fluent and adequate, and whether the expected simplifications were performed. These aspects were judged in a scale of **1 (worse)** to **3 (best)**. For simplification quality, in particular, the 1-3 scores were given according to the following criteria: 1 = none of the expected simplifications was performed or at least one inadequate simplification was performed; 2 = some of the expected simplifications were performed, but not all of them; and 3 = all of the expected and possibly additional adequate simplifications were performed. The average scores obtained for the 20 segments are: **Fluency = 2.5**, **Adequacy = 2.5**; **Simplification =**

2.35. Some examples of original segments (source), their simplified version according to our model (target) and the human annotator (reference) are shown in Table 4.

Table 4. Examples of source, simplified and reference segments

Source	Target	Reference
no dia da liberdade de impostos, criado pela associação da classe média (aclame) *para marcar a* data em que os brasileiros, simbolicamente, deixam de trabalhar para pagar tributos, *foram vendidos* em novo hamburgo gasolina, computadores e um carro.	no dia da liberdade de impostos, criado a associação da classe média (aclame), *com o objetivo de marcar* a data em que os brasileiros, simbolicamente, deixam de trabalhar para pagar tributos. *em novo hamburgo, nesse dia, venderam-se* gasolina, computadores e um carro.	a associação da classe média (aclame) criou o dia da liberdade de impostos. *o objetivo é marcar* a data em que os brasileiros, simbolicamente, deixam de trabalhar para pagar tributos. *em novo hamburgo, nesse dia, venderam-se* gasolina, computadores e um carro.
segundo o presidente da comissão de finanças da assembléia, ricardo sanguino, o processo ...	*de acordo com* o presidente da comissão de finanças da assembléia, ricardo sanguino, o processo ...	ricardo sanguino é o presidente da comissão de finanças da assembléia. *ele disse que* o processo ...
o anúncio do resultado ocorre em um megaevento *em lisboa, portugal, amanhã à noite*.	o anúncio do resultado ocorre em um megaevento *em lisboa, amanhã à noite em portugal* .	o anúncio do resultado ocorre *amanhã à noite em um grande evento em lisboa. lisboa fica em portugal* .
segundo bush, as mudanças em anbar são um exemplo ...	*na opinião de* bush, as mudanças em anbar são um exemplo ...	*na opinião de bush*, as mudanças em anbar são um exemplo ...
o animal deu alguns passos e começou a *verter sangue pela boca.*	o animal deu alguns passos e começou a *verter sangue pelo estabelecimento.*	o animal deu alguns passos e começou a *despejar sangue pela boca*.
-- eu sei que é *imprudência, mas* só dirijo na praia, *só para* aprender.	-- eu sei que é *descuido. mas* só dirijo na praia*, para* aprender.	-- eu sei que é *descuido. mas* só dirijo na praia*, para* aprender.
o relatório *produzido pela* comissão criada para revisar processos administrativos ...	o relatório *da comissão* criada para revisar processos administrativos ...	o relatório *da comissão* criada para examinar os processos administrativos ...
ontem, a moeda norte-americana fechou cotada a r 2,026*, queda* de 3,20%	*ontem*, a moeda americana fechou cotada a r 2,026. *a queda foi* de 3,20%	a moeda norte-americano fechou a r 2,026 *ontem. a queda foi* de 3,20%
tombou primeiro sob as duas patas da frente *e então* caiu de vez .	tombou primeiro sob as duas patas da frente. *em seguida* caiu de vez.	ele tombou primeiro sob as duas patas da frente*. em seguida,* caiu de vez .
mesmo que a incerteza sobre o tamanho da crise *dificulte* projeções de prazo mais longo*, seus* efeitos abrem janelas para ...	*a incerteza* sobre o tamanho da crise *dificulte* projeções de prazo mais longo. *mas seus* efeitos abrem janelas ...	*a incerteza* sobre o tamanho da crise *torna difíceis* as projeções de prazo mais longo. *mas seus* efeitos abrem janelas para ...

According to the manual inspection, as we can see by the examples in Table 4, the lexical operations (complex words and discourse markers) are in general well captured. Some short-distance reordering is also correctly performed, mostly for

adverbs. Simple cases of sentence splitting are performed correctly, such as that of clauses separated by conjunctions. Long distance operations like canonical (S-V-O) ordering and clause inversion in general are not performed. The most problematic cases are very long sentences, particularly those containing brackets: the model usually tries to remove the brackets, but cannot always deal with their content adequately. Some of the deficiencies are due to the nature of the approach, since only information about words and short phrases is used.

4 Conclusions and future work

We have presented a new approach for Text Simplification, based on the framework of Statistical Machine Translation. The experiments have shown that the framework can appropriately simplify certain phenomena, particularly those related to lexical operations: lexical simplification and simple rewriting. As for the other operations, the system is usually overcautious and does not perform simplifications. While this yields low recall, it guarantees precision, by preventing complex operations to be applied when they would most likely degrade the results.

These experimental results in general are very promising, particularly considering the size of the parallel corpus and the fact that a basic SMT system was used as such, without any adaptation to the simplification task. It is well-known that SMT requires considerably more data than the amount used in this paper. For training standard systems to translate news texts like ours, it is recommended a few hundred thousands sentence pairs, but we only used ~4,000 sentence pairs. While translating between variations of the same language alleviates the need for such large corpora, the fact that several types of simplification were considered here makes the task considerably harder. Finally, complex operations like clause reordering depend on deeper information than words and phrases. Certain additional features, like part-of-speech tags, can already be used as *factors* in standard phrase-based systems like Moses, although more training data would be necessary. The use of other information sources, particularly syntax, is also possible with frameworks for hierarchical SMT [20].

We believe that having a larger corpus of appropriately annotated original-simplified texts, covering enough examples of the several types of simplifications, is very likely to yield a larger number of "simple" simplifications without degrading the quality of the outcome segments. With a larger parallel corpus, one could build one "translation" model for each type of simplification and then apply all models to a given complex input text, either in sequence or in combination.

As future work, besides building larger corpora, we plan to investigate how the standard SMT framework could be adapted to further improve the results. We also consider adding syntactic information to the framework in order to better address long-distance operations, like clause inversion. Finally, we would like to evaluate whether the simplifications produced here are useful for a given target user by testing them within a certain task, such as answering a questionnaire or solving some problem.

References

1. Ribeiro, V. M.: Analfabetismo e alfabetismo funcional no Brasil. In: Boletim INAF. Instituto Paulo Montenegro, São Paulo (2006)
2. Max, A.: Writing for Language-impaired Readers. In: Proceedings of 7th Conference on Intelligent Text Processing and Computational Linguistics, Mexico City (2006) 567-570
3. Petersen, S. E.: Natural Language Processing Tools for Reading Level Assessment and Text Simplification for Bilingual Education. PhD thesis, University of Washington (2007)
4. Siddharthan, A.: Syntactic Simplification and Text Cohesion. PhD thesis, University of Cambridge (2003)
5. Devlin, S., Unthank, G.: Helping aphasic people process online information. In: Proceedings of the ACM Conference on Computers and Accessibility, Portland, Oregon (2006) 225-226
6. Klebanov, B., Knight, K., Marcu, D.: Text Simplification for Information-Seeking Applications. LNCS 3290, Springer-Verlag, New York (2004) 735-747
7. Vickrey, D., Koller, D.: Sentence Simplification for Semantic Role Labeling. In: Proceedings of the ACL-HLT (2008) 344-352
8. Chandrasekar, R., Srinivas, B.: Automatic Induction of Rules for Text Simplification. Knowledge-Based Systems, 10 (1997) 183-190
9. Daelemans, W., Hothker, A., Sang, E. T. K.: Automatic Sentence Simplification for Subtitling in Dutch and English. In: Proceedings of the 4th Conference on Language Resources and Evaluation, Lisbon, Portugal (2004) 1045-1048
10. Petersen, S. E., Ostendorf, M.: Text Simplification for Language Learners: A Corpus Analysis. In: Proceedings of the Speech and Language Technology for Education Workshop, Pennsylvania, USA (2007) 69-72
11. Candido Jr., A., Maziero, E., Gasperin, C., Pardo, T.A.S., Specia,L., Aluisio, S.M.: Supporting the Adaptation of Texts for Poor Literacy Readers: a Text Simplification Editor for Brazilian Portuguese. In: Proceedings of the NAACL/HLT Workshop on Innovative Use of NLP for Building Educational Applications, Boulder, Colorado (2009) 34–42
12. Gasperin, C., Specia, L., Pereira, T., Aluisio, S.M. Learning When to Simplify Sentences for Natural Text Simplification. In: Proceedings of the Encontro Nacional de Inteligência Artificial (ENIA), Bento Gonçalves, Brazil (2009) 809-818
13. Simard, W., Goutte, C., Isabelle, P. Statistical Phrase-based Post-editing. In: Proceedings of NAACL HLT, Rochester, USA (2007) 508–515
14. Caseli, H.M., Pereira, T.F., Specia, L., Pardo, T.A.S., Gasperin, C., Aluísio, S.M. Building a Brazilian Portuguese parallel corpus of original and simplified texts. In: 10th Conference on Intelligent Text Processing and Computational Linguistics, Mexico City (2009) 59-70
15. Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, C., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., Herbst, E.. Moses: Open Source Toolkit for Statistical Machine Translation. In: Proceedings of the 45th ACL, demonstration session, Prague, Czech Republic (2007)
16. Och, F. J. Minimum error rate training in statistical machine translation. In: Proceedings of the 41st ACL. Sapporo, Japan (2003) 160-167
17. Papineni, K., S. Roukos, T. Ward, and W. Zhu. Bleu: a method for automatic evaluation of machine translation. In: Proceedings of the 40th ACL, Morristown (2002) 311-318
18. Doddington, G. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In: Proceedings of the 2nd Conference on Human Language Technology Research, San Diego (2002) 138–145
19. Callison-Burch, C., Koehn, P., Monz, C. and Schroeder, J. Findings of the 2009 Workshop on Statistical Machine Translation. In: Proceedings of the 4th Workshop on Statistical Machine Translation, Athens, Greece (2009) 1-28.
20. Chiang, D. A hierarchical phrase-based model for statistical machine translation. In Proceedings of the 43rd ACL, Ann Arbor, USA (2005) 263-270