Karlsruhe Institute of Technology

Error analysis of a statistical machine translation system

Bachelor's Thesis of

Sabrine Chelbi

at the Department of Informatics
Interactive Systems Lab (ISL)
Institute for Anthropomatics and Robotics

Reviewer:	Prof. Dr. Alexander Waibel
Second reviewer:	Dr. Sebastian Stüker
Advisor:	Dr. Jan Niehues

Duration: 20 Januar 2016 – 19 Mai 2016

Karlsruher Institut für Technologie
Fakultät für Informatik
Postfach 6980
76128 Karlsruhe

I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

Karlsruhe, 19.05.2016

. . . . . .

(Sabrine Chelbi)

Abstract

Machine Translation (MT) is nowadays widespread and has gained tremendous momentum. It allows to a large number of persons to participate in the information revolution due to its fluent application in different fields. With no human involvement, MT performs the translation of a text from a source language into a target language using computers.

Different approaches were introduced to solve the problem of natural language understanding like the rule-based MT (RBMT) and statistical MT (SMT). The RBMT translates the text using different rules. These rules are developed manually by human experts. In contrast to the RBMT approach, the statistical MT generates its translations using statistical models. The parameters of these models are obtained from the analysis of bilingual text corpus. It generates multiple solutions and chooses the best fluently and grammatically correct sentences. Most of the MT systems use the statistical model. It leads to a better quality of the system.

Nowadays, Machine Translation has a broad application in different domains. In order to evaluate the MT, automatic and manual approaches can be employed. Various automatic methods has been proposed to measure the performance of the machine translation output like Bleu and Meteor. Manual evaluation can be also employed and gives a better result in order to measure the quality of the MT and to analyze the errors of the system output.

In this project, we use an inhouse machine translation system, which is developed by the Interactive Systems Labs (ISL). The data held in this thesis is taken from newspapers, Ted and TedX talks. It is a set of bilingual translations : from German to English and from English to German. We try to carry out a detailed error analysis of the machine translation system in order to enhance the MT output. Therefore, we apply the Part-Of-Speech tagging on the available reference and candidate translations to have an idea about the differences on the number of tags between both translations. We have also generated Out-Of-Vocabulary (OOV) words of each language pair translation. The goal of this step is to make an estimation about the number of words that our system could not translate. Afterwards, we compute the BLEU score of each language pair to concentrate on the problematic translations.

In order to focus more on the generated errors, we calculate the BLEU scores of each sentence of the system translations. Afterwards, a manual analysis is applied to analyse the errors and to find the systematic problems.

Zusammenfassung

Die Maschinelle Übersetzung (MÜ) ist sein langer Zeit eine interessante Forschungsbereich mit vielen Innovationen. Sie wird heutzutage in vielen Bereichen eingesetzt. Ohne menschliche Beteiligung, führt die MÜ die Übersetzung eines Texts aus einer Ausgangssprache in eine Zielsprache unter Einsatz von Computern.

Um das Problem der natürlicher Sprachverarbeitung zu lösen, wurden verschiedene Ansätze wie regelbasierte MÜ (RBMÜ) und statistiche MÜ (SMÜ) eingeführt. Die MÜ übersetzt den Text mittels unterschiedlichen Regeln. Diese Regeln werden manuell von menschliche Experten entwickelt. Im Gegensatz zu der RBMÜ, erzeugt die SMÜ seine Übersetzungen durch statistische Modelle. Die Parameter dieser Modelle werden aus der Analyse des bilingualen Korpora erhalten. Es erzeugt mehrere Lösungen und wählt die besten fließend und grammatisch korrekte Sätze. Das statistische Modell ist am meisten benutzt. Es führt zu einer besseren Qualität des Maschinelle Übersetzungssystems.

Heutzutage, die MÜ hat eine breite Anwendung in verschiedenen Bereichen. Um die MÜ zu evaluieren, werden automatische und manuelle Methoden eingesetzt. Verschiedene automatische Methoden wurden vorgestellt wie Bleu und Meteor um die Übersetzungsqualität zu bewerten. Um die Fehler der Systemausgabe zu analysieren, können auch manuelle Methoden eingesetzt werden.

In diesem Projekt, verwenden wir ein Inhouse-Maschinenübersetzungssystem , das von den Interactive Systems Labs (ISL) entwickelt wird. Die in dieser Arbeit verwendeten Daten sind aus Zeitungen, Ted und TEDx Gespräche genommen. Sie enthalten zweisprachigen Übersetzungen: vom Deutschen ins Englische und vom Englischen ins Deutsche. Wir versuchen, eine detaillierte Fehleranalyse des maschinellen Übersetzungssystems auszuführen, um die MÜ Ausgabe zu verbessern. Deshalb, verwenden wir die Part-Of-Speech tagging auf den verfügbaren Referenz and Kandidaten Übersetzungen um den Unterschied zwischen die beiden Übersetzungen zu erläutern. Wir haben auch Out-of-Vocabulary (OOV) Wörter jedes Sprachenpaar Übersetzung erzeugt. Das Ziel dieses Schrittes ist eine Abschätzung über die Anzahl der Wörter, die das System nicht übersetzen kann, zu bilden. Danach berechnen wir die Bleuscore jedes Sprachenpaar um auf die problematischen Übersetzungen zu konzentrieren.

Um mehr über die generierten Fehler zu konzentrieren, berechnen wir die BLEU-Werte jeden Satze der System Übersetzungen. Danach wird eine manuelle Analyse ausgeführt um die Fehler zu analysieren und die systematischen Probleme zu finden.

Contents

1. Introduction

This an introductory chapter. Section 1.1 presents a motivation for the problem handled in this thesis. The goals of this thesis are discussed in Section 1.2. Section 1.3 describes the structure of this thesis.

1.1. Motivation

The production of a machine translation system is a very hard problem. Languages are complex and for each language it is necessary to take the world knowledge into consideration. Many words have different meanings and can be differently translated. The error analysis of a machine translation system presents a difficult problem. The evaluation of the machine translation output can be done automatic. Different metrics have been proposed like Bleu, WER,METEOR. Only automatic evaluation of the translation output is not sufficient to determine the source of errors. That is why, a manual analysis is needed to identify the main problems of a machine translation output. The determination of the main errors is important to focus the effors on rectifying these errors in order to improve the machine translation system.

The goal of this thesis is to analyse the errors of a statiscal machine translation. The machine translation system is developed by the Interactive System Labs (ISL) in the Karlsruhe Institut of Technology (KIT). The error analysis in this work will be based on an automatic and a manual analysis of the output. The data used are collected from Ted and TedX talks and newspapers. In this thesis, we aim to provide a detailed error analysis in order to enhance the inhouse statistical machine translation system.

1.2. Goals of the project

In this project, we aim to make a detailed error analysis of our inhouse statistical machine translation system. Our goal is to find the errors which should be rectified to improve the machine translation system. In order to achieve that, we make a detailed analysis of the translations by comparing for example the BLEU score of the reference and system translations or the number of verbs, nouns,etc using the POS tagging. In addition, we make a manual analysis of the translation output and as a result we can cite the errors that the system committed and that can be adressed to enhance the system.

1.3. Structure of the thesis

After the introduction in the Chapter 1, the related work is presented in Chapter 2. Chapter 3 introduces the foundations which are important to understand this thesis. We describe the approach used in this thesis in Chapter 4. Then, we present the data, which is used in this work, in Chapter 5. Chapter 6 discusses the methods used to evaluate the data and presents the experimental results analyzing the errors of the system. Finally, Chapter 7 includes a summary of this work and an outline of the future work.

2. Related Work

Different automatic evaluation for machine translation were introduced like BLEU and METEOR. A framework to analyse the translation errors automatically is presented by [24]. It is based on the use of morpho-syntactic information. The goal is to have details about the nature of translation errors. The error analysis is carried out on an English-Spanish statistical machine translation system. The morphological inflections are eliminated from the refernce and the system translations. WER and PER are calculated on distinct word classes and forms. The results obtained by this method are similar to the results obtained by a human error analysis, but not the same.

Using only the automatic evaluation is not sufficient to analyze the errors of a MT system. Thus multiple manual evaluation for MT has been introduced.

[27] presents an open source tool for error analysis of a Machine Translation output. It is a graphical user interface and facilitates the annotation of the translations in order to identify the similarities between them. It can be used with any Machine Translation system and between any language pairs. In addition to that, any error topology can be employed.

Different taxonomies for MT error analysis were introduced. The most used is the error topology introduced by [28]. The proposed error classification has a hierarchical structure and is divided into five big classes: "Missing Words", "Word Order", "Incorrect Words", "Unknown words" and "Punctuation" errors. It has been applied by different others researchers like [14] and [25] in order to enhance the performance of statistical machine translation. [28] has presented a framework for the error classification of a MT system. An error analysis of the system was performed. It was used by the RWTH in the first TC-STAR evaluation. The language pairs held in this paper are English-Spanish and Chinese-English translations. The European Parliament Plenary Sessions (EPPS) corpora is used for the English-Spanish translation and the broadcast news for the Chinese-English language pair. A detailed analysis of the results is presented. The most important error for the English-Spanish translation is the incorrect generation of the verb tense. For Chinese-English translation, the most occured error is the word order.

In this work, we will perform the proposed error classification on our English-to-German and German-to-English translation to obtain a detailed analysis of our in-house system.

Different erros can occur in a MT and make the system worse. One of these errors is the reordering problem. That is why different reordering models were introduced to enhance the machine translation systems. In a recent work, [17] developed a structure-aware reordering approach which was applied in a German-English machine translation system. The approach consists of comparing the translation outputs of two different translation systems, which applied POS and syntax trees reordering rules. As a result, enhancing the position of the verb or translating verbs that could not be translated before exhibits an observable improvement.

In order to evaluate a machine translation quality, [19] has performed an error analysis on a machine translation system. The goal is to find criteria to assess the translation quality concerning the accuracy of semantic content in translations. It has been demonstrated how an error analysis can be a starting point for criteria to evaluate the MT quality. The error classification described in this paper concentrates on mismatching semantic components in the source and target texts in contrast to [28], which compares the reference and system translations. It evaluates differing patterns between human and machine translations. In addition, it reveals the differences between two types of machine translation systems. In the statistical MT system, the most common error is the incomprehensible "word salad". In the rule-based MT system, the mistranslating is the most occured error and which leads to misleading sentences. Thus, the errors occured in the machine translation systems are related to the relations between the concepts. In the human translation, it is important to differentiate true amissions from implications and true additions from explication to evaluate the semantic quality.

3. Basics

3.1. Machine Translation

Machine translation is a translation of a text from one language to another with no human involvement using computer software. It exists three basic approaches of a machine translation:

1. transfer-based MT: This approach is based on morphological, syntactic and semantic analysis of the source and target language. The RBMT uses different rules which are developed manually by human experts.

2. interlingual MT: The source language is transformed to a language-independent representation which is an abstract description of the semantics of the text.

3. direct MT: Source language is translated to the target language directly without any intermediate stage in the translation process. The morphological inflection are removed from the input to obtain the base form which will be searched in a bilingual dictionary.

The main difference between these three approaches is the depth of the analysis of the source language. The Vauquois Triangle 3.1 [16] [1] modularizes the process into steps. For each step up, the approach requires more analysis on the source language side and more generation on the target language side. The direct MT, which represents the level at the bottom in the triangle, is a direct lexical conversion of the source language to the target language. The top level represents the Interlingua where each sentence is transformed into an abstract representation.

The statistical MT [22] requires a large bilingual text corpora to generate translations using statistical models. This approach generates multiple solutions and chooses the best fluently and grammatically correct sentences. The SMT uses probabilistic models to find the most likely target sentence ẽ for a source sentence f. Using the Bayes' theroem, the best translation can be described as follow:

$$\tilde{e} = \arg\max_{e} P(e|f) = \arg\max_{e} P(e)P(f|e)$$

P(e) is the language model which measures the fluency of the proposed output sentence. P(f|e) represents the translation model which indicates a set of possible translations with their probabilities for the source sentence. The best translation ẽ will be found by selecting the sentence with the highest probability.

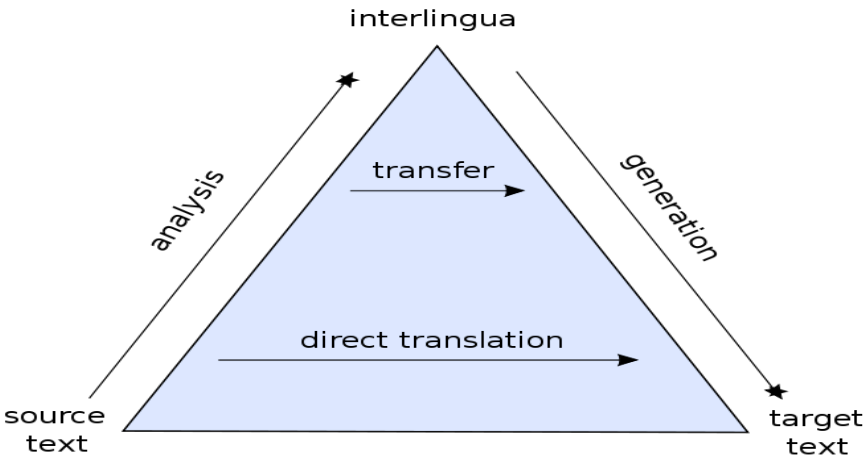

Figure 3.1.: The Vauquois-Triangle

3.2. Manual evaluation

An automatic evaluation gives a single quantitative score, which measures the performance of the machine translation, but it does not give any information about the weaknesses of the system. A manual evaluation allows the system developers to measure the quality of their MT and to analyze the errors of the system output.

Errors can be classified as a hierarchical structure 3.2 as described in [28]. The classification of the errors is splitted into five big classes which are: "missing words", "word order", "incorrect words", "unknown words" and "punctuation". A "missing word" error can be caused when a word is lacked. We can distinguish two types of errors: the first when missing words leads to a meaningless sentence like nouns, verbs,etc. The second type is when missing words is only necessary to build a gramatically correct sentence.

The next class of type errors is "word order", which is splitted in two categories: "word-level" and "phrase-level" reordering and within each of these categories we can distinguish between local or long range reordering. If a "word order reordering" error occurates, then individual words should be moved to generate a correct sentence. In the case of "phrase level reordering", a block of consecutive words should be reordered. The difference between local and long range is that words in the first type should be reordered in a local chunk unlike the "long range", where words should be reordered into another context.

The third error category is the "incorrect words". It is caused when the system does not find a correct translation of a word. We can distinguish five subcategories: "Sense" which injures meaningless sentence, "incorrect form" of a word, "extra words" in the target sentence, "style" which means a bad choice of words and "idioms" which

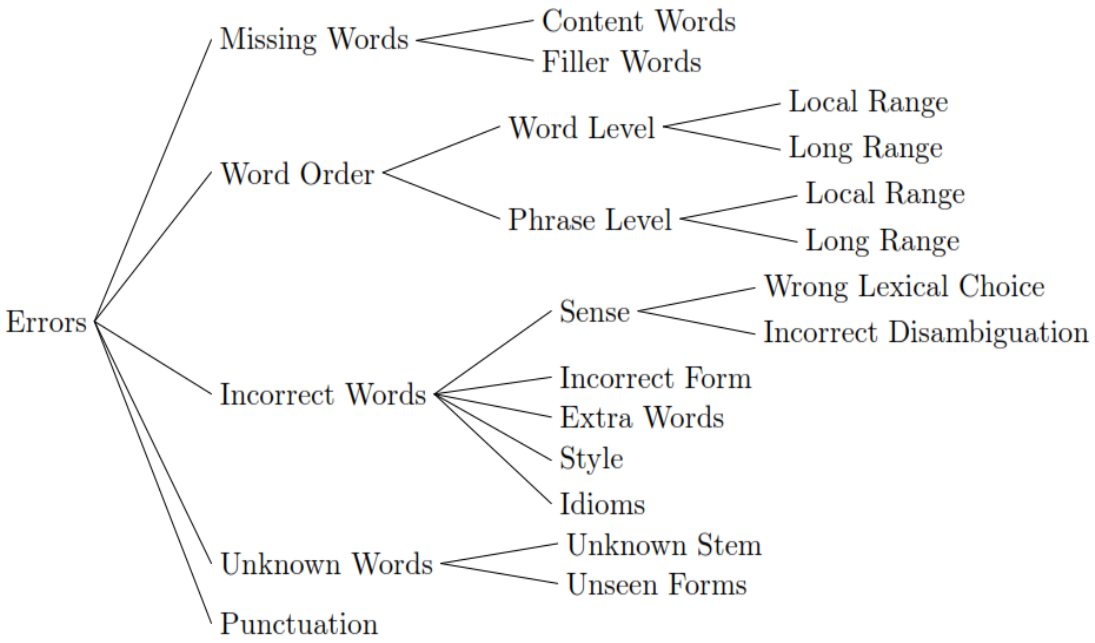

Figure 3.2.: Classification of translation errors[28]

are idiomatic expressions that the system tries to translate as a normal text.

Last but not least, the "unknown words" errors which are splitted into "truly unkown words" and "unseen words".

The last error type is the "punctuation", which represents minor disturbances for the current machine translation output.

3.3. Automatic evaluation

Human evaluation is one among different methods of machine translation evaluation. It is extensive, but it is very expensive, needs a lot of human resources and can take months to finish. That is why an automatic evaluation is needed. It should be inexpensive, quick, consistent and correlate well with the human evaluation. Various automatic evaluation methods for machine translation have been employed such METEOR,Bleu, Word Error Rate,etc. The most used method is the Bleu method.

3.3.1. Bleu

Bleu is an abbreviation for "Bilingual Evaluation Understudy" and it is introduced by [23]. It is an alogrithm for an automatic evaluation of the quality of a machine translation. This method compares the hypothesis translation to one or more reference translations. The hypothesis translation, also candidate translation, is a translation

that has to be evaluated. The reference translation is usually the correct translation, at which the hypothesis will be compared. The translation is good, when the candidate translation shares many words and sentences with the reference translations. The Bleu's formula for calculating the score is:

$$Bleu = BP * exp(\sum_{n=1}^{4} w_{\mathrm{n}} * \log p_{\mathrm{n}})$$

The computation of the Bleu-Score is based on a modified n-gram precision. We calculate the number of words of the system translation which occur in any single reference translation and then we divide by the total number of words in the candidate translation. The algorithm computes for each sentence the 1-, 2-, 3- and 4-grams. The BP (Brevity Penalty) is used when the hypothesis translation is shorter than the reference translation. The Bleu-Score is always between 0 and 1. This score illustrates how much the system and the reference are similar. When the value is closer to one, then the system translation have a lot of similar sentences with the reference translation and thus the translation is good.

3.3.2. TER

TER is an abbreviation for "Translation Error Rate". It is an automatic evaluation metric of a machine translation output. It was proposed by [26]. It measures the number of edits required that a human would have to perform to change a system output into one of the references. The edits needed can be insertions, deletions, substitutions of words or shifts of a word sequence. TER is defined as :

$$TER = \frac{\text{number of edits needed}}{\text{average number of reference words}} \tag{3.1}$$

Consider the reference and system sentences below: Reference: "Jetzt kannst du auf eine richtige Schule gehen," sagte er. System: "Du kannst jetzt zu einer echten Schule gehen", sagte er. The reference and system sentences have the same meaning, but TER considers that they are not the same because it does not exist an exact match between the two sentences. First of all, "Du" and "jetzt" should be removed. In addition, the words "zu einer echten" must be substitued with "auf eine richtige". Totally, the number of edits is 4: 3 substitutions and 2 shifts. The reference sentence contains 10 words. Thus, the TER score is calculated as :

$$TER = \frac{4}{10} = 40 \tag{3.2}$$

3.3.3. METEOR

METEOR [18] is a metric for automatic evaluation of machine translation. It is an improvement of the BLEU method and is based on precision and recall. METEOR [20] uses explicit word-to-word matching between the system and the reference translation. If there is multiple reference translations, every hypothesis translation is

matched with the reference translation and the best score will be used. The algorithm creates first a word alignment between the system and the reference translation. The alignment is a set of matchings between words. If there exists two alignments with the same number of mappings, then the alignment which has fewer intersections, is chosen. Every word in the hypothesis sentence is mapped to at most one word in the reference sentence. The algorithm uses three modules of word matching. First of all, the "exact module" is applied, where two identical words are matched. Then the "Porter Stem module" is used. The unigrams are stemmed using the "Porter Stemmer" algorithm. This algorithm removes the morphological and inflexional suffix from words in English. After suffix stripping, two words are matched if they are the same. Finally, the "Synonymy module" is called. Two words are matched, if they are synonym. If the words belong to the same group of cognitive synonyms in "WordNet", then they are considered synonym. The "WordNet" is a database for english words that also exists for other languages. After these three modules, we obtain a final alignment and the METEOR score can be calculated.

$$score = (1 - Penalty) * F_{\text{mean}} \tag{3.3}$$

The harmonic mean F_{mean} is calculated as follow :

$$F_{\text{mean}} = \frac{10PR}{R + 9P} \tag{3.4}$$

The unigram recall R is declared as below:

$$R = \frac{m}{w_{\text{r}}} \tag{3.5}$$

m is the number of common words in the system and reference translation. w_r is the number of words in the reference translation. P is a unigram precision :

$$P = \frac{m}{w_{\text{t}}} \tag{3.6}$$

m is defined above. w_t is the number of words in the system translation. So the harmonic mean F_{mean} is calculated by combining precision and recall based on unigram matches. To take into consideration larger common segments in reference and hypothesis translation, a penalty P is computed using large n-gram mapping.

$$Penalty = 0.5 * (\frac{c}{u_{\text{m}}})^3 \tag{3.7}$$

c is the number of chuncks. Unigrams are classified into a few possible number of chuncks. A chunck is a group of adjacent words in the system and reference translation. u_m is the number of mapped words.

3.4. Part-Of-Speech

The Part-of-speech [2] is a category of words that have a common grammatical features. The words are classified according to their significance, their morphological

form or their syntactic usage in the sentence. The POS tags that are frequently used are verbs, adverbs, nouns, adjectives, conjunctions, cardinal numbers,etc.

The main goal of using the POS information is to disambiguate different morphological categories. As an example, we have this sentence : "Many talk the talk".The word "talk" can be a noun which refers to a conversation or a verb which means communicating by spoken words..

The POS Tagging [3] is the process of assigning the POS to each word in the text. There are two distinctive groups of POS tagging algorithms: rule-based and stochastic tagging.[4]

1. Rule-based Tagging [15] :

   This tagging is based on two steps:

   - Fist of all, to each word are applied all the possible POS tags. For example, we have this phrase: "His talk is about his boss". The word "talk" is tagged as a verb and a noun.
   - In the second step of the algorithm, some rules must be used to determine the correct POS tag. As an example for the rule: "If the keyword is preceded by a possessive pronoun, then it is a noun and not a verb". For the previous example, the word “talk” will be tagged as a noun and not a verb.

2. Stochastic Tagging [5]:

   This kind of tagging is based on statistical models. We suppose that each word has a finite set of possible tags. These tags can be drawn from a dictionary or a morphological analysis. If a word has many possible tags, statistical methods enable us to determine the optimal sequence of POS tags.

In order to assign POS tag to each word, a POS taggger is needed. It is a tool that reads a text and annotates words with part-of-speech. Stanford-Log-Linear POS Tagger [6] is used in this thesis. It is a stochastic tagging and is a Java implementation of the Log-linear POS tagger developed by the Stanford Natural Language Processing Group [7]. The POS tagger contains multiple models like English, German, French, etc. The Tag set is a list of tags that can be assigned to a word. It is different depending on the language.

For English, the most known tag set and which is used in this thesis is called the "Penn Treebank Tag set" [8]. The Figure 3.3 shows an example of alphabetical list of POS tags used. For example, an adjective will be assigned as a "JJ" tag. A singular noun will be annotate as a "NN" tag. The complete alphabetical list of part-of-speech

Number	Tag	Description
1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential *there*
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending

Figure 3.3.: Example of Penn Treebank POS tags [9]

tags used in the Penn Treebank Project is available online [9].

As an example 3.4, we take a sentence from the German-EnglishTed translation. We apply then the "Penn Treebank Tag set" so that a tag will be assigned to each word.

```
Sentence without POS Tags: When I was 11 , I remember waking up one morning to the sound
of joy in my house .
Sentence wit POS Tags : When_WRB I_PRP was_VBD 11_CD ,_, I_PRP remember_VBP waking_VBG
up_RP one_CD morning_NN to_TO the_DT sound_NN of_IN joy_NN
in_IN my_PRP$ house_NN ._.
```

Figure 3.4.: Example of applying POS Tags

As we can see in the Figure 3.4, a "CD" tag is assigned to the cardinal number "11". The word and its POS tag are attached with an underscore.

The German taggers use the “Stuttgart-Tübingen Tag set” [10]. In the Figure 3.5, you can find a sublist of STTS tags with their definitions and examples. For example, an adverb will be annotated with a "ADV" tag. A "CARD" tag will be assigned to each cardinal number.

POS	DESCRIPTION	EXAMPLES
ADJA	attributives Adjektiv	[das] große [Haus]
ADJD	adverbiales oder prädikatives Adjektiv	[er fährt] schnell, [er ist] schnell
ADV	Adverb	schon, bald, doch
APPR	Präposition; Zirkumposition links	in [der Stadt], ohne [mich]
APPRART	Präposition mit Artikel	im [Haus], zur [Sache]
APPO	Postposition	[ihm] zufolge, [der Sache] wegen
APZR	Zirkumposition rechts	[von jetzt] an
ART	bestimmter oder unbestimmter Artikel	der, die, das, ein, eine
CARD	Kardinalzahl	zwei [Männer], [im Jahre] 1994
FM	Fremdsprachliches Material	[Er hat das mit ``] A big fish [" übersetzt]

Figure 3.5.: Example of Stuttgart-Tübingen tag set [10]

4. Approach

The goal of this thesis is to make a detailed automatic and manual analysis of our machine translation system in order to enhance the system output.

This chapter describes the approach that we will employ to assess the MT system.

In order to find the frequent errors, we will try to cluster the errors generated by the SMT. First of all, we will cluster errors by using the POS tags, because the system might do the same errors to the same POS tags. We will apply the Part-Of-Speech Tagging on the reference and system translations of each language pair. The POS tagging is employed on the translations in order to annotate each word with the following POS tags such as verbs, nouns, etc. After applying the POS tagger on the reference and the system translation of each language pair, we get the text so that a POS tag is assigned to each word. Afterwards, we compute the number of occurrences of each tag and save it in a comma-separated values (csv) file. The csv file contains multiple lines, where each line is composed from a POS tag and its value for the reference and the system translation separated by a comma. In order to plot these csv files to histograms, the matplotlib [11] is used. It is a 2D plotting library for Python that allows to generate mathematical representations. The pyplot [12] interface provides a MATLAB-like plotting framework and is available for simple plotting. That is why, the matplotlib.pyplot interface is used in this thesis. As a result, we will obtain a histogramm for each language pair that contains the number of POS tags in the reference and the system translations. We can compare the two translations and have an idea about the type of the missing words.

Next to the missing words problem, another typical problem in the error analysis of a SMT is the Out-Of-Vocabulary (OOV). That is why the OOV words will be generated for each language pair. OOVs are words that exist in the test data but not in the traningsdata. The goal of this step is to have an idea about the number of OOVs and what type of words occures in the OOV list. The percentage of the OOVs of each language pair will be presented in Chapter 6.

Moreover, in order to analyze automatically the missing words and OOV errors, the Bleuscores of each Language Pair (LP) translation will be computed using an inhouse script. The scores are saved in a csv file. This file is plotted using the matplotlib.pyplot interface described above to a histogram, which contains the Bleuscores of all translations. This step aims to find the most problematic LP translations in order to more concentrate on these LP on the manual analysis.

To identify the source of errors and to identify the most prominent errors in the translations is not possible. That is why we will try to calculate the Bleuscores for each sentence of the LP translations. The scores are computed using an inhouse script and are sorted in a descending order. The 20 lowest scores are selected. Their conrresponding source, reference and system sentences are generated and are saved in text files.

To focus more on the generated errors, we will analyze the 20 obtained sentences from the previous step manually, beacuse it is not possible to make a manual analyse of all the data. We will compare the reference and the system translations and try to generate the errors in a hierarchical structure as described in the Chapter 3.2. The figure 4.1 resumes the approach described above.

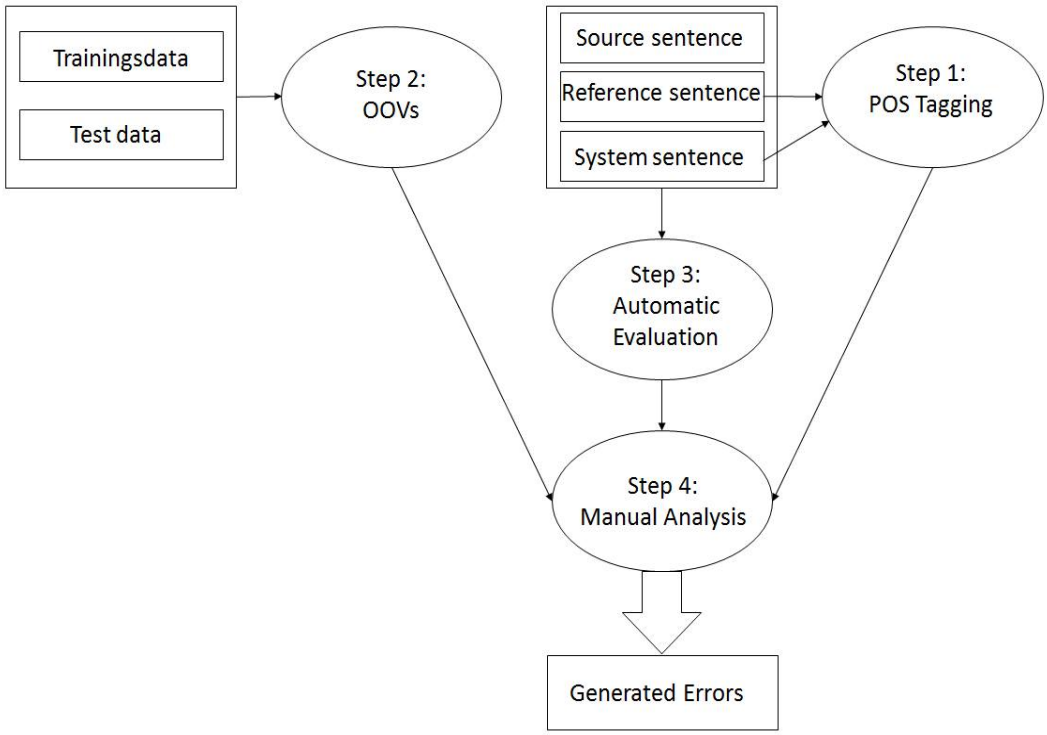

Figure 4.1.: Approach

5. Data

The error analysis held in this thesis is performed on an inhouse decoder system output. The data is a set of bilingual translations : from English to German and from German to English. For each translation, it exists three files :

- source file : It contains data that will be translated
- reference file : It is the human translation, to which the system translation will be compared.
- system file : It represents the translation, that the used system generates.

In order to generate these translations, two systems were used : one system for the WMT evaluation and one system for the IWSLT evaluation.

The WMT system was employed to the German-EnglishNews translation. For this translation, the data are articles, which are taken from different newspaper. The source file, which is in german, contains 3003 sentences, a total of 68545 words. The english reference data is composed also from 3003 sentences, 59325 words totally. The data, that are generated by th system, contains 64563 words.

The IWSLT system was employed to the German-EnglishTed, German-EnglishTedX and English-GermanTed translations.

For the German-EnglishTed and English-GermanTed translation, the data are talks taken from Ted conferences. The Technology Entertainment and Design is a set of international conferences, which cover different topics in more than 100 languages. Ted is organized by the non-profit "Sapling Foundation" and its goal is resumed in its slogan : diffuse "Ideas worth spreading". The speakers should present their ideas in an innovative way in form of short talks of maximum 18 minutes. The best conferences are available gratis.

The German-EnglishTed source file contains 993 sentences, a total of 16687 words. The english reference data is composed also from 17811 words. The data, that are generated by th system, contains 20459 words.

Concerning The English-GermanTed translation, the source file contains 66102 words, the reference and the system file have 993 sentences : 16687 words in the reference file and 19731 words in the sysem file.

For the German-EnglishTedX translation, the talks are taken from TedX conferences. TedX is an event like Ted. It is planned by independent organizers in their own country. The EnglishTedX source file contains 1363 sentences. The reference file is composed from 23346 words and the system file from 23304 words.

5.1. WMT System

The WMT System is a phrase-based SMT system and is presented in the 7th Workshop on statistical Machine Translation WMT12. The system uses a discriminative word alignment method as described in [21]. It uses the conditional random field CRF to model the alignment matrix and to add features easily. CRF is an unidirectional probabilistic graphical model that models the conditional distribution randomly. The algorithms, that are used in the CRFs, are modified in order to model dependecies between different random variables. The language model used in this system is 4-gram SRI language model. This LM employs the Kneser-Ney Smoothing, which is trained by the SRILM toolkit. The SRILM is an extensible language modelin toolkit. It is a set of C++ libraries and scripts in order to enable creation and evaluation of different N-gram LM. The translation system uses the POS-based and tree-based packages to model the problem of word reordering. To obtain the POS tags, the TreeTagger model is employed. The translation models are trained on the EPPS and News Commentary Corpora. This system is tuned with the news-test2011 data.

5.2. IWSLT System

The system is presented in the 12th International Workshop on Spoken Language Translation. It is a phrase based translation system. It uses an inhouse phrase-based decoder described in [29].

First of all, the system makes the preprocessing of the data. Then GIZA++[1] Toolkit is used to make the word alignments over the data. The language model used in this system is 4-gram LM with modified Kneser-Ney Smoothing, which is trained by the SRILM toolkit. A word-based and contextual-based language model are used. A 9-gram sequence of POS tags are utilized. In the translation system, two types of word reordering are employed.

In the 1st type, some reordering rules are applied on the source data and each resulting source sentence is encoded in a lattice.

In the 2nd reordering type, each phrase pair is scored from the phrase table and the word alignment. Then, a lexicalized model accumulates the reordering probabilities.

6. Experiments and results

This chapter surveys the various experiements applied on the data of this thesis and an error analysis of the system will be done in order to improve our machine translation system.

6.1. Detailed analysis of the data

6.1.1. POS tagging

As described in the Chapter 4, the first applied step is the POS tagging. After applying the POS tagger and computing the number of occurrences of the POS tags, we obtain four histograms in total. For each language pair, there is a histogram comparing the occurrence of POS tags of the reference and the system translation.

Figure 6.1 shows an example of histogram of the English-GermanTed translation.We can see that there is a difference on the number of occurrences of verbs and symbols. Symbols present 17.31% in the system translation and in the reference 15.13%. Symbols can be punctuations like comma or other marks like [,),'.

The most remarkable difference is the occurrence of the verbs, which is more notable in the reference translation with 16,79% than in the candidate translation with 14,91%. The main reason of lacking verbs can be that some verbs are not considered in the system translation. When a sentence contains more than one verb, it can be that one verb is missed in the system translation. The lack verbs reasons will be more analyzed in details in the manual error analysis. As we can see in the histogram, there is no big differences in the other POS tags.

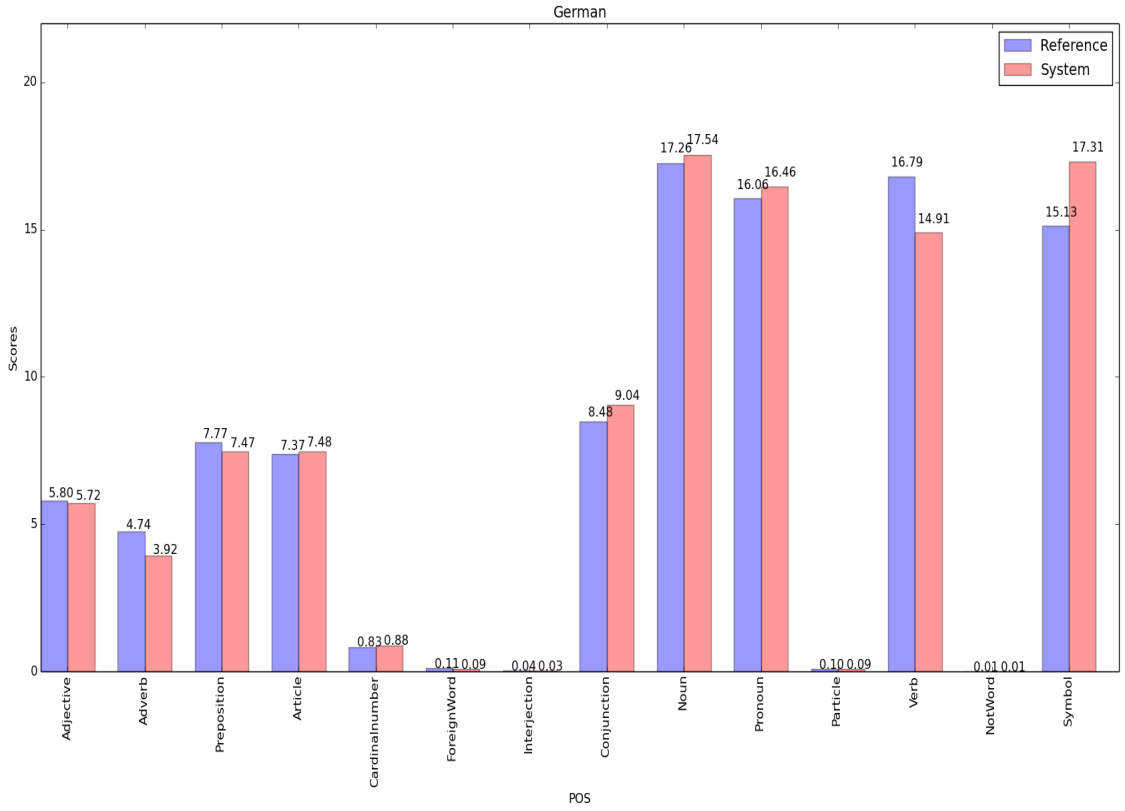

Figure 6.1.: occurrence of POS tags of the English-GermanTed translation

The next figure 6.2 presents the histogram of the occurrence of POS tags of the German-EnglishTedX. We note also an observable difference in the occurrence of adverbs. It exists more adverbs in the system translation 9,56% than in the reference 7,66%. The source sentences are speechs taken from the TedX and it can be that the speaker uses a lot of adverbs. The reference is translated using less adverbs in constrast to the source. That is why in the system translation we find more adverbs than in the reference. This can be a cause to have more adverbs in the system translation. In the manual analysis, we will focus on this error.

We note also a litte variance in the verb tags. In the reference translation, it exists 15,54% of verbs more than in the system translation 14,35%. Long and complicated sentences can be a cause of missing verbs in the system translation. This will be manually analyzed in the next chapter. As we can observe in the histogram, there is no big differences in the other POS tags like pronouns, nouns, adjectives.

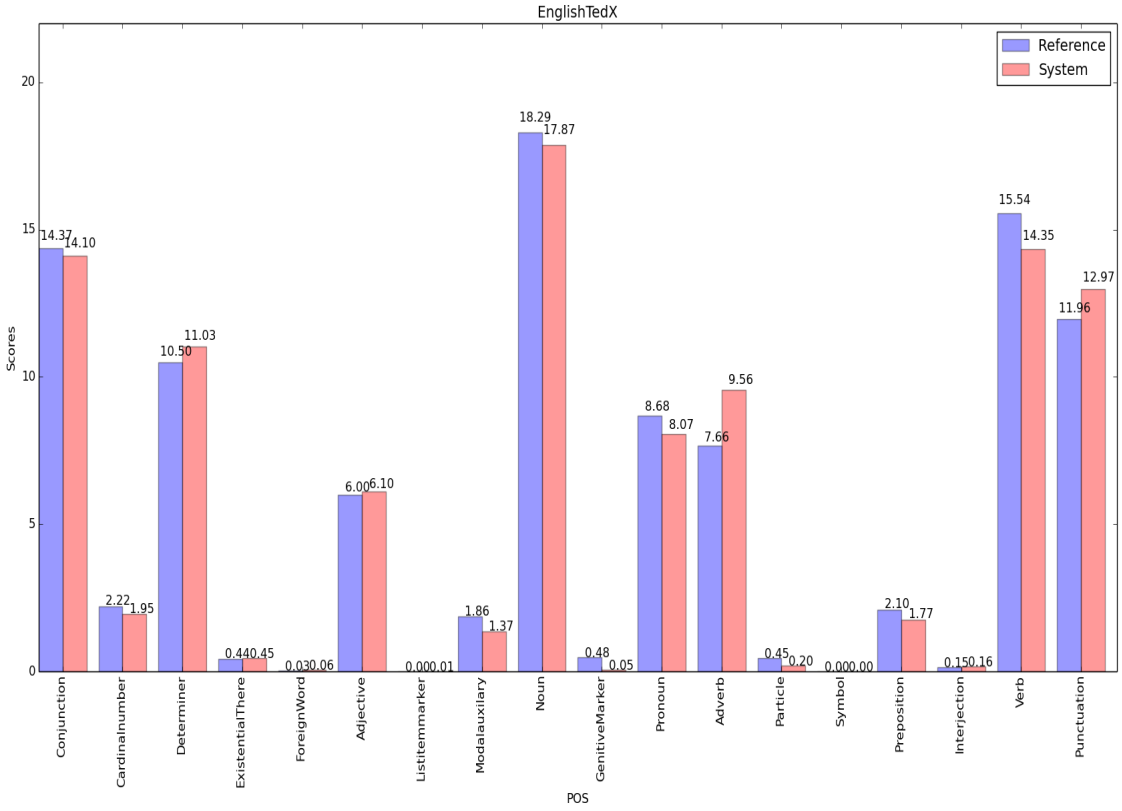

Figure 6.2.: occurrence of POS tags of the German-EnglishTedX translation

For the German-EnglishNews histogram 6.3, there a difference in the number of occured verbs. In the system translation, we note 12,18% and in the reference translation the verbs present 14,53%. A reordering problem can cause missing verbs error . This lacking of verbs will be more analyzed manually.

We can also notice that it exists more determiners in the system translation with a variance of 2%. Determiners can be articles like "no", "the" and indefinite determiners like "another", "neither". It can be that the system translates a plural noun word by word and the reference translates it without determiner. For example, in the source sentence we have the words "die Beinen". The system translates thes words in "the legs", but in the reference sentence we find only "legs". This can be a cause of having more determiners in the system translation. The number of conjunctions is more in the candidate translation with 15,33% than in the reference with 13,80%. It can be that the system sentence contains a lot of determiners or conjunctions and the system tries to translate the source sentence word by word but the reference is translated in another way. This can be a reason of having more determiners and conjunctions in the system translation than in the reference one. We will concentrate more on these errors in the manual analysis.

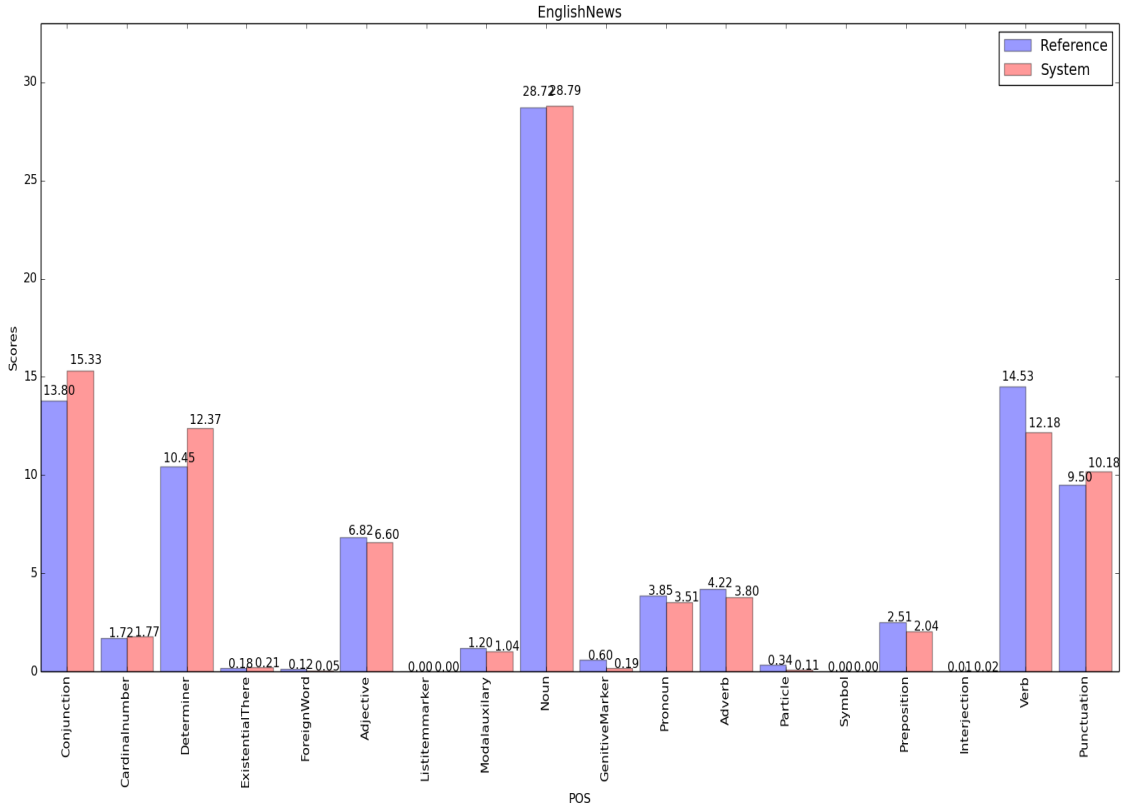

Figure 6.3.: occurrence of POS tags of the German-EnglishNews translation

For the German-EnglishTed language pair 6.4., we note a difference in the number of verbs with a variance of about 1,3%. Verbs are missed in the system translation because the system tries to translate the source sentence word by word. It may that the reference sentence contains more verbs than the source and then the system sentence will contain more verbs than the reference. The lacking of verbs will be analyzed in the manual evaluation.

We remark also that there is a distinction in the occurence of the nouns: in the reference translation 18,58% and in the candidate translation 19,14%. The system translation contains more nouns than the reference one and thus can refer that both translations are correct but each one is translated in a different way. We will explain it more in the manual analysis.

As we cann see in the histogram, there is no big differences between the system and the reference translation in the number of occurrence of the other POS tags.

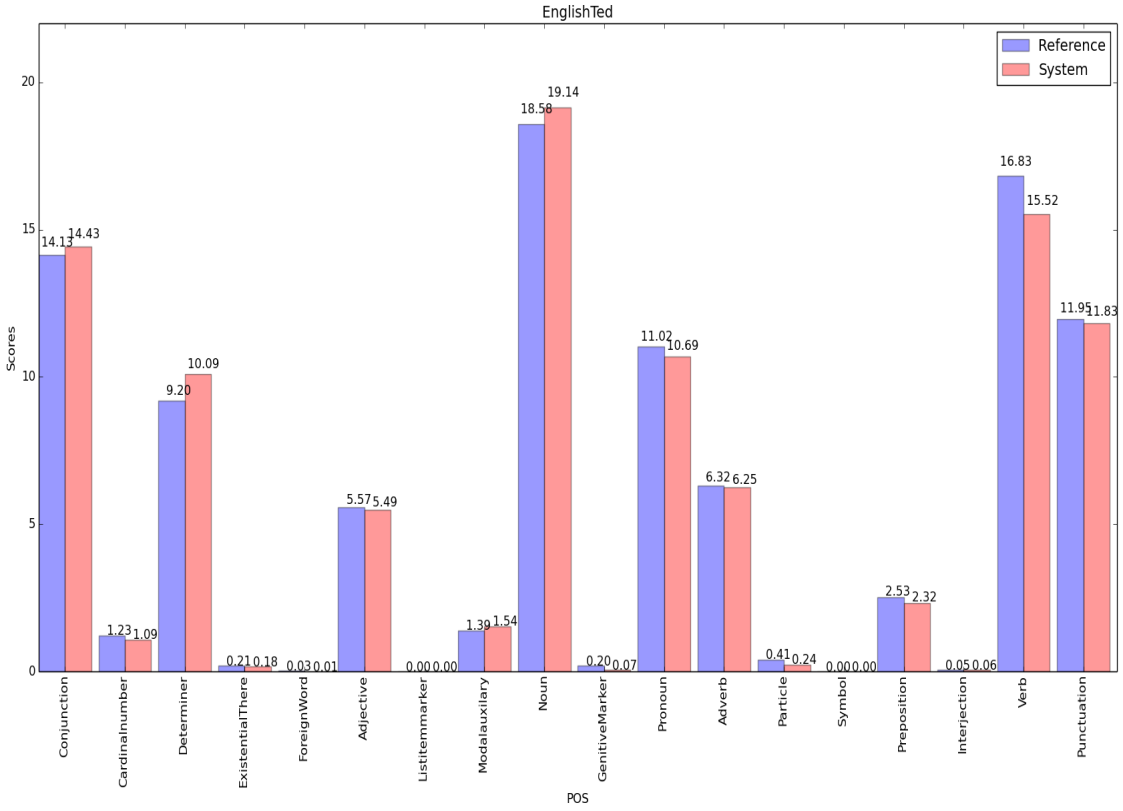

Figure 6.4.: occurrence of POS tags of the German-EnglishTed translation

In order to analyze in further details the error reasons, the errors are examined more deeply in the manual analysis.

6.1.2. OOV

Out-of-vocabulary words are words that appear in the candidate translation but not in the recognition vocabulary [13]. Thus, they are words that are unknown for the system. So, they could not be translated.

In this thesis, OOVs are generated for each language pair translation and are saved in files. the OOVs are attached in the appendix. For the English-GermanTed, there is 43 OOVs. It represents a percentage of 0,22%. The 43 found OOVs are nouns with "ing" at the end, words that begin mit uppercase, proper nouns,etc. Nouns ending with "ing" are stemmed from verbs, which exist already in the vocabulary. We can remark that some of the OOVs are due to writing errors in the input.

For the German-EnglishNews, we have found 667 OOVs which is 1,03% of the hole number of words. Compound nouns, website name or numbers are examples of the not translated words. The german language contains multiple compound nouns and should be splitted so that the system can translate these words. Our system does not

understand the composition of these nouns, that is why they are considered as OOV words.

The German-EnglishTed has 88 OOVs, which represents only 0,43%, and the German-EnglishTedX has 96 OOVs, which represents 0,41%. As an example of OOVs, there are separable verbs and compound nouns. Another comming OOV in the German-EnglishTedX is the colloquial words, that can not be translated.

As we can remark, there is no large problem of OOV words.

6.1.3. Bleu Score

To enhance our inhouse MT system and to identify the errors, we analyze the MT system by computing the Bleuscore for each language pair translation using an in-house script. The goal of this step is to have an idea about the existing errors in order to focus on in the manual analysis.

As we can see in the figure 6.5, the English-GermanTed translation has the worstest Bleuscore 26,21%. A possible reason can be that the source text is complicated. Afterwards comes the German-EnglishNews with 27,77%. The German-EnglishTedX has the score 28,69%. Finally, the German-EnglishTed has the best score 33,98%.

So as to concentrate more in the details and detect the errors, we decide to compute the Bleuscore for each sentence using an inhouse script. The obtained scores are sorted upward. Then, we select the 20 sentences, which have the worstest Bleuscores. The Bleuscore of the worsest sentences for the English-GermanTed varies between 2.47 and 4.11, for the German-EnglishNews between 1.06 and 3.27, for the German-EnglishTed between 2.94 and 5.52, and finally for the German-EnglishTedX between 2.56 and 4.58. These generated sentences will be manually analyzed in order to detect the errors that cause awful scores.

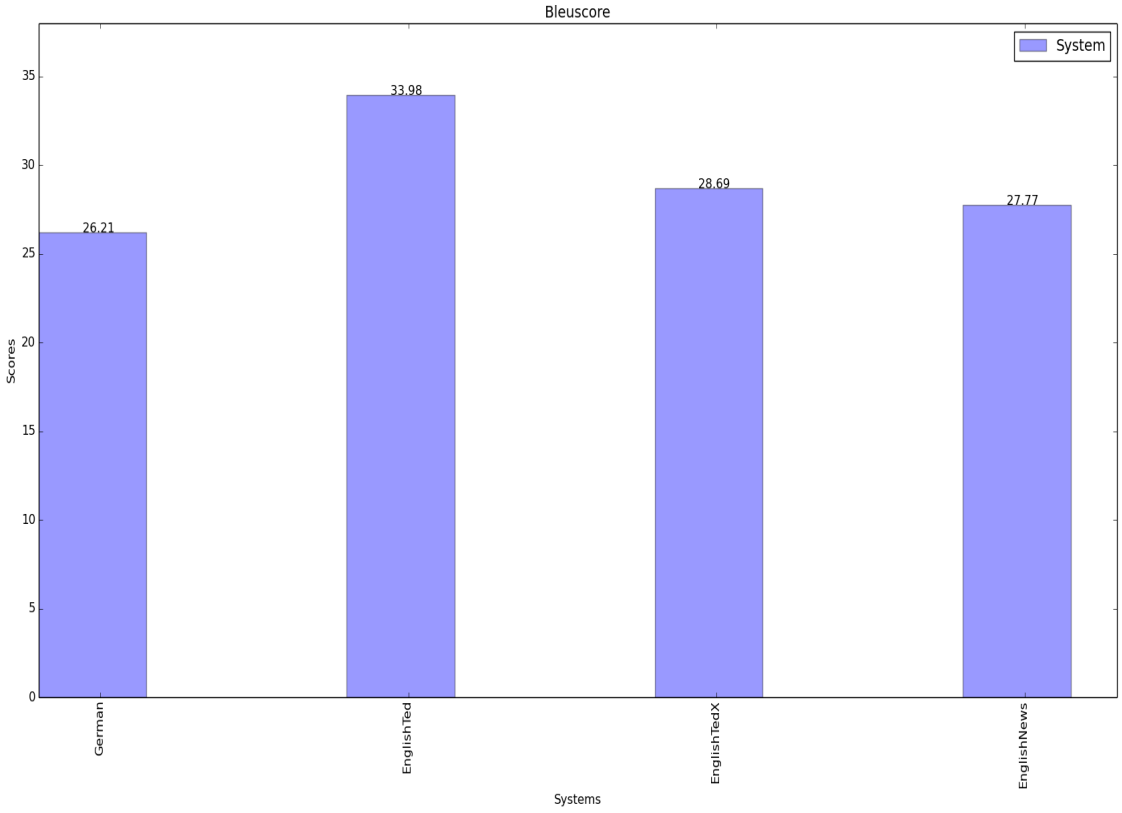

Figure 6.5.: Bleuscore of the different language pair translations

6.2. Error analysis

In the previous section, we made some experiments to have an idea on which type of errors should we more concentrate in the manual analysis.

Our manual analysis is applied on the 20 worstest sentences of each language pair translation, which have been generated with the help of the Bleuscore. The reference and system translation are compared and analyzed and the errors are obtained in a hierarchical structure as described in the chapter 3.2. The German-EnglishNews and English-GermanTed translations have the worstest Bleuscores. That is why we will concentrate our analysis more in this language pairs.

6.2.1. German-EnglishNews translation

The most occurred error is missing words. Missing words can be content or filler words. 13 sentences suffer from this error and the causes are different. The reasons of missing words in these 13 worstest sentences are described below.

1. The first reason of missing words is that the reference translation contains sometimes more contents than the source sentence. Usually, it exists words that have their explanations in parentheses. We can cite this example:

   Source sentence: um 19 Uhr gestaltet die biraböhmische Blasmusik aus Schömberg den großen Unterhaltung Abend.
   Reference sentence: At 19:00, the "biraböhmische" (a play on words alluding to the regions pear trees, Bohemia, and perhaps their love of beer as well!) wind music group from Schömberg, provide a great evening of entertainment.
   System sentence: At 7.00 p.m. the biraböhmische brass bands from Schömberg large entertainment evening .

   In the above example, the adjective "biraböhmische" is explained in the reference translation. Thus, it leads to a missing word error.

2. Verbs are sometimes missed, which leads to a missing content word error and thus the sentence will be meaningless. We remark, that this error occur mostly when the sentence is more complicated or long, i.e. it is not a normal sentence. In addition to that, reordering problem can cause missing verbs. It exists sentences, that have the same structure but one of them the verb is missed and the other sentences contain the verbs and are correct. We can cite this example:

   Source sentence: der Text vermittelt sich auf diese angestrengte Weise jedoch kaum .
   Reference sentence: However, the source text makes barely any reference to this intense delivery.
   System sentence: The text on the hard way .

3. Idiomatic expressions can lead to a meaningless sentence, because the system tries to translate these expressions word by word and then chooses the incorrect words. You can the following example:

Source sentence: sie nannte Sprich Wörter wie : Holz Auge sei wachsam , wie Schuppen von den Augen fallen , ein Auge auf jemand werfen , den Seinen gibt es der Herr im Schlaf , seine Hände in Unschuld waschen .
Reference sentence: She referred to sayings such as: "Holzauge sei wachsam" (keep your eyes peeled), "wie Schuppen von den Augen fallen" (like scales falling from one's eyes), "ein Auge auf jemand werfen" (to cast an eye on someone), "den Seinen gibt's der Herr im Schlaf" (good things come to some when they sleep) and "seine Hände in Unschuld waschen" (to wash one's hands of something).
System sentence: She called sayings like : wooden eye is vigilant , as to the fact that an eye on someone , the gang , there is the Lord in his sleep , wash his hands .

4. An incorrect word error happens when the conversion of a unit measure is not done. We can cite this example :

Source sentence: Fahrer , der mit 210 km / h und heißem Getränk zwischen den Beinen gerast war , erhält 1000 £ Strafe
Reference sentence: Driver speeding at 130mph with hot drink between legs fined £1,000
System sentence: Driver , who was with 210 km / h , and hot drinks between the legs gerast gets £ 1000 fine

5. A rarely occuring error is the word ordering error. When a sentence is long or complicated, the system tries to translate the source sentence word by word or the reference translation is translated with another way. Thus, a word order error happens as we can see in this example.

Source sentence: Passagiere können auch ausschließlich beim Direkt Kauf auf der Website von Frontier Plätze reservieren .
Reference sentence: It also allows passengers to choose their seat in advance only if they buy directly from the Frontier website.
System sentence: Passengers can also book exclusively to the direct purchase on the website of Frontier .

6. In this 20 worstest generated sentences, we note that unknown words are found like: "Zollernalbkreis", which is a compound noun, "Australierin", which means a woman which has the australian natinality, and "gerast" the participle II of the verb "rasen".

   Another causes can injure multiple errors but it is not engendered by the system. We remark, that the system translation is sometimes better than the reference translation. It could happen that the system translation conforms more to the source sentence than the reference sentence as this example.

Source sentence: NSA erklärt , ein " interner Fehler " und nicht Hacker seien Schuld am Crash der Website
Reference sentence: NSA Blames "Internal Error," Not Hackers, For Website Crash
System sentence: NSA declared that an " internal error " and not hackers are to blame for the crash of the website

7. In the system translation, we find sometimes extra words like determiners or conjunctions. This occurs because the system tries to translate the source sentence word by word and the reference sentence is translated in a different way than the system sentence. For example, the word "die Beinen" is translated by th system in "the legs" and in the reference is written only the word "legs". That is why sometimes we find more words in the system than en the reference translations.

6.2.2. German-EnglishTed and German-EnglishTedX translation

The most occured problem for both language pair translations is that most of the source sentences are taken out of the context. They are sometimes not completed sentences as we can remark in this German-EnglishTed example:

Source sentence: König nennen , und ihn womöglich noch auf Ideen zu seiner Nachfolge bringen .
Reference sentence: They didn't want to call him King in case that gave him ideas, or his successor ideas.
System sentence: King , and perhaps even to bring ideas to his successor .

We can cite an example from the German-EnglishTedX translation too:

Source sentence: die schwimmen halt und sollen dann an ihren ja Gelenken dann über Hydraulik den Strom produzieren .
Reference sentence: They swim and are then supposed to produce electricity in their , well , limbs , via hydraulic power .
System sentence: The stop and then swim to their joints then yes on hydraulics produce the electricity .

Thus, this leads to a bad choice of words and incorrect words. Source sentences play a key role in the translations because not completed sentences or incorrect sentences will lead to big problems in the system translations.

Moreover, it happens rarely in the German-EnglishTed translation that the system sentence is better than the reference sentence.This was an example:

Source sentence: aber das stimmt nicht . es geht zurück auf einen Befürworter der Presse Freiheit .
Reference sentence: It's not. It's down to a campaigner for the freedom of the press.
System sentence: But that is not true . It goes back to an advocate of press freedom .

In addition to that, sometimes the reference sentence is translated in a different way than the system sentence and it contains more verbs than the source. Consequently, the candidate translation will contain less verbs than the reference translation. This explains the lacking of verbs in the system translation. We can cite this example:

Source sentence: manche werden es vielleicht nicht bis zum Ende schaffen .
Reference sentence: You see, some people may not do it. They may not get through it.
System sentence: Some are perhaps not going to make it to the end .

We have remarked in the German-EnglishTedX translation that the adverbs are missed in the system translation. After analyzing the source sentences, we observe that the speaker uses many adverbs and the reference is translated in a different way so that it does not contain the same number of adverbs. The system tries to translate the source sentence word by word that is why the system translations contain more adverbs than the reference translation. We can cite this example:

Source sentence: so sagen es jeden falls auch die Kalküls .
Reference sentence: That ' s what those who are calculating the costs are saying , at any rate .
System sentence: Anyway , so it also say the calculations .

6.2.3. English-GermanTed translation

After the manual analysis of the sentences, we have remarked that the negation is not considered in the system translation, which causes an alignment problem. As an example the "'t" in the source sentence is not translated in the system translation. Thus a missing word error is caused and the translated sentence will have another sense. For example:

Source sentence: we don 't bump into every neighbor , so a lot of wisdom never gets passed on , though we do share the same public spaces .
Reference sentence: Wir treffen nicht jeden Nachbarn auf der Straße, sodass viele Weisheiten nicht weitergegeben werden. Doch wir nutzen dieselben öffentlichen Plätze.
System sentence: Wir stoßen in jeder Nachbar , so viel Weisheit wird nie weitergegeben , obwohl wir die gleiche öffentliche Räume teilen .

In addition, sometimes the second verb of the sentence is not translated when the sentence contains more than a verb. This causes a missing word error and a meaningless sentence too. As an example, we can cite:

Source sentence: I will put down the lawsuit and they will send me all the information I ask for .
Reference sentence: Ich ziehe die Anklage zurück, wofür sie mir alle verlangten Informationen schicken.
System sentence: Ich werde die Klage , und sie werden mir all die Informationen , die ich fragen .

Moreover, a word order error happens rarely when, for example, it exists a conjunction and the verb must be in the end of the sentence. An example is given below. However the system sentence is not correct, but the word ordering error can be noted

after the conjunction "weil" in the translation sentence generated by the system:

Source sentence: you already tried it once , right , and you %-% I tried it before , but I stopped because it gave me a shock .
Reference sentence: Du hast es bereits einmal versucht, nicht wahr – Ja, ich hab's schon mal versucht, aber den Versuch wieder aufgegeben, weil ich ... einen Schlag bekam.
System sentence: Sie versuchte es schon einmal , und ich versuchte es , aber ich hörte auf , weil es gab mir einen Schock .

Furthermore, in the english language it exists expressions that means that the noun is in plural like "set of rules" which cann be translated in German "Regeln". The problem of the system that it tries to translate these expressions word by word. Consequently, extra words are generated and the meaning of the sentence is not absolutely correct. For example:

Source sentence: you don 't have to design a whole new set of hospitals to do it .
Reference sentence: Man muss keine neuen Kliniken dafür bauen.
System sentence: Sie müssen eine neue Reihe von Krankenhäusern , es zu tun .

We would note that these analysis is based only on the 20 sentences, which have the worstest Bleuscore. It may be exist other errors that we have not mention here, because there were not exposed in these sentences.

In the above automatic analysis, we have got an idea about the occurred errors in the different language pair translations. We have analyzed these translations manually and got almost the same errors as in the automatic evaluation. We remark an agreement between both evalution methods.

7. Summary and Future Work

In this project, we try to make a detailed error analysis of the statistical machine translation system developed by the Interactive Systems Labs (ISl). Our goal is to indicate the errors which should be settled in order to enhance our machine translation output. Therefore, we have applied different experiments on the data used in this thesis. Fist of all, we have employed Part-of-Speech (POS) Tagging on the different reference and candidate language pair translations. Secondly, Out-Of-Vocabulary (OOV) are generated for each translation. Afterwards, Bleuscore of each language pair translation is computed. Moreover, Bleuscore of each sentence is calculated and the 20 sentences of each language pair, which have the worstest Bleuscores, are selected. Based on these generated sentences, we carried out our manual analysis in order to detect the errors caused by the system. We have begun our analysis by the German-EnglishNews translation where the most occured errors are missing words. We have remarked, that sometimes the reference translation has more content than the system translation. Other errors can occur even if they are not engendered by the system. We can cite as an example that sometimes the system translation is more better than the reference. In the German-EnglishTed and German-EnglishTedX, the most occured error in this thesis is that most sentences are not completed and taken out of context. This causes different errors like missing words, incorrect words and causes meaningless sentences. Finally in the English-GermanTed translation, the most arised error is missing content words like verbs. Sometimes sentences suffer from word ordering error. In addition to that, sometimes source sentences contain incorrect characters or errors. This causes the appearence of errors in the system translation.

As part of our future work is to focus more on the choice of the source sentences mostly of the German-EnglishTed and German-EnglishTedX translation. To make a detailed error analysis of the system translation, we should have correct data. So the source sentences should be completed sentences and not taken of context. That is why we consider that it is interessant to concentrate on the choice of source sentences.

A further objective is to concentrate on the Out-Of-Vocabulary (OOV) problem. It can be reduced by using the stemming algorithm. For example, in the German-EnglishTed we have an OOV word, which is "sociopolitically", even though the word "sociopolitical" exists in the phrase table. In the English-German translation, we note different OOVs, which are mostly seperable verbs. The number of OOVs, that can be reduced by stemming, is very small but we consider that is interessant to focus on this problem.

Bibliography

[1] In: URL: http://www.translationdirectory.com/articles/article1945.php.

[2] In: URL: http://www.computational-logic.org/iccl/master/lectures/summer06/nlp/part-of-speech-tagging.pdf.

[3] In: URL: http://language.worldofcomputing.net/pos-tagging/parts-of-speech-tagging.html.

[4] In: URL: https://user.phil-fak.uni-duesseldorf.de/~petersen/Einf_CL/Part-of-Speech-Tagging.pdf.

[5] In: URL: http://ccl.pku.edu.cn/doubtfire/NLP/Lexical_Analysis/Word_Segmentation_Tagging/POS_Tagging_Overview/POS%20Tagging%20Overview.htm.

[6] In: URL: http://nlp.stanford.edu/software/tagger.shtml.

[7] In: URL: http://nlp.stanford.edu/.

[8] In: URL: http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/Penn-Treebank-Tagset.pdf.

[9] In: URL: https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html.

[10] In: URL: http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-table.html.

[11] In: URL: http://matplotlib.org/.

[12] In: URL: http://matplotlib.org/users/pyplot_tutorial.html.

[13] In: URL: https://www.lti.cs.cmu.edu/sites/default/files/research/thesis/2013/long_qin_learning_out-of-vocabulary_words_in_automatic_speech_recognition.pdf.

[14] Eleftherios Avramidis and Philipp Koehn. "Enriching Morphologically Poor Languages for Statistical Machine Translation". In: *ACL 2008, Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics, June 15-20, 2008, Columbus, Ohio, USA*. 2008, pp. 763–770. URL: http://www.aclweb.org/anthology/P08-1087.

[15] Eric Brill. "A Simple Rule-Based Part of Speech Tagger". In: *Proceedings of the Third Conference on Applied Natural Language Processing*. Trento, Italy: Association for Computational Linguistics, Mar. 1992, pp. 152–155. URL: http://www.aclweb.org/anthology/A92-1021.

[16] Bonnie J. Dorr, Eduard H. Hovy, and Lori S. Levin. "Machine Translation: Interlingual Methods". In: 2006.

[17] Teresa Herrmann, Jan Niehues, and Alex Waibel. "Manual Analysis of Structurally Informed Reordering in German-English Machine Translation". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), Reykjavik, Iceland, May 26-31, 2014*. 2014, pp. 4379–4386. URL: http://www.lrec-conf.org/proceedings/lrec2014/summaries/569.html.

[18] P. Koehn. "Statiscal Machine Translation". In: 2010, pp. 228–229.

[19] Maarit Koponen. "Assessing Machine Translation Quality with Error Analysis". In: URL: http://www.sktl.fi/@Bin/40701/Koponen_MikaEL2010.pdf.

[20] Alon Lavie and Michael J. Denkowski. "The Meteor Metric for Automatic Evaluation of Machine Translation". In: vol. 23. 2-3. Sept. 2009, pp. 105–115.

[21] Jan Niehues and Stephan Vogel. "Discriminative Word Alignment via Alignment Matrix Modeling". In: *Proceedings of the Third Workshop on Statistical Machine Translation*. 2008, pp. 18–25.

[22] Miles Osborne. "Statistical Machine Translation". In: *Encyclopedia of Machine Learning*. 2010, pp. 912–915. URL: http://dx.doi.org/10.1007/978-0-387-30164-8_783.

[23] Kishore Papineni et al. "BLEU: A Method for Automatic Evaluation of Machine Translation". In: *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. ACL '02. Philadelphia, Pennsylvania: Association for Computational Linguistics, 2002, pp. 311–318. DOI: 10.3115/1073083.1073135. URL: http://dx.doi.org/10.3115/1073083.1073135.

[24] Maja Popoví et al. "Morpho-syntactic Information for Automatic Error Analysis of Statistical Machine Translation Output". In: *Human Language Technology Conf. / North American Chapter of the Assoc. for Computational Linguistics Annual Meeting (HLT-NAACL), Workshop on Statistical Machine Translation*. New York City, NY, USA: Association for Computational Linguistics, June 2006, pp. 1–6.

[25] Maja Popovic and Aljoscha Burchardt. "From human to automatic error classification for machine translation output". In: *15th International Conference of the European Association for Machine Translation (EAMT 11). Annual Conference of the European Association for Machine Translation (EAMT-11), 15th, May 30-31, Leuven, Belgium*. European Association for Machine Translation, May 2011. URL: http://www.dfki.de/web/research/publications/renameFileForDownload?filename=Popovic-14-finalVersion.pdf&file_id=uploads_1121.

[26] Matthew Snover et al. "A study of translation edit rate with targeted human annotation". In: *In Proceedings of Association for Machine Translation in the Americas*. 2006, pp. 223–231.

[27] Sara Stymne. "Blast: A Tool for Error Analysis of Machine Translation Output". In: *Proceedings of the ACL-HLT 2011 System Demonstrations*. Portland, Oregon: Association for Computational Linguistics, June 2011, pp. 56–61. URL: http://www.aclweb.org/anthology/P11-4010.

[28] D. Vilar et al. "Error Analysis of Statistical Machine Translation Output". In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)*. ACL Anthology Identifier: L06-1244. Genoa, Italy: European Language Resources Association (ELRA), May 2006. URL: http://www.lrec-conf.org/proceedings/lrec2006/pdf/413_pdf.pdf.

[29] Stephan Vogel. "SMT DECODER DISSECTED : WORD REORDERING". In: 2003.

A. Appendix

A.1. Sentences, that have the worstest Bleuscore in the English-GermanTed translation

Source: we don 't bump into every neighbor , so a lot of wisdom never gets passed on , though we do share the same public spaces .

Reference: Wir treffen nicht jeden Nachbarn auf der Straße, sodass viele Weisheiten nicht weitergegeben werden. Doch wir nutzen dieselben öffentlichen Plätze.

System: Wir stoßen in jeder Nachbar , so viel Weisheit wird nie weitergegeben , obwohl wir die gleiche öffentliche Räume teilen .

Source: so what that means is that people will be able to — People will be able to call in from their mobile phones and do this test , and people with Parkinson 's could call in , record their voice , and then their doctor can check up on their progress , see where they are doing in this course of the disease .

Reference: Das heißt, die Leute können – Die Leute können mit dem Handy anrufen und den Test machen. Leute mit Parkinson könnten anrufen, ihre Stimme aufnehmen lassen, sodass ihr Arzt den Fortschritt der Krankheit überprüfen kann.

System: Das bedeutet , dass Menschen in der Lage sein werden , People - wird in der Lage , aus ihren Mobiltelefonen und diesen Test , und die Menschen mit Parkinson nennen könnten , in ihrer Stimme , und dann ihren Arzt auf ihren Fortschritt sehen , wo sie in diesem Kurs der Krankheit .

Source: the morans are the warriors who protect our community and the livestock , and they are also upset about this problem .

Reference: Unsere Krieger heißen Morans. Sie beschützen unseren Stamm und unsere Herden. Sie sind ebenfalls aufgebracht wegen dieses Problems

System: Die morans sind die Krieger , die unsere Gemeinschaft und Vieh schützen , und sie sind auch aufgeregt über dieses Problem .

Source: the House of Representatives didn 't want Washington to get drunk on power .

Reference: Das Repräsentantenhaus wollte nicht, dass Washington machtgierig wird. Sie wollten ihn nicht

System: Die Abgeordnetenkammer wollte Washington betrunken .

Source: we delight at ballet virtuosos and tap dancers you will see later on .

Reference: Wir sind entzückt beim Anblick von meisterhaften Ballett- und Stepptänzern, wie Sie gleich sehen werden.

System: Wir bei Ballett Virtuosen und Tänzer sehen Sie später .

Source: they didn 't want to call him King in case that gave him ideas , or his successor ideas .

Reference: König nennen, und ihn womöglich noch auf Ideen zu seiner Nachfolge bringen.

System: Sie wollen nicht nennen ihn König für den Fall , dass ihm Ideen , oder sein Nachfolger Ideen .

Source: now , that would actually save on a difficult trip to the clinic , and what if patients could do that test themselves , right ?

Reference: Das würde eine beschwerliche Tour ins Krankenhaus ersparen. Was wäre, wenn Patienten diesen Test selbst machen könnten?

System: Nun , das würde auf eine schwierige Reise in die Klinik zu retten , und was , wenn die Patienten tun könnten , die sich selbst testen , nicht wahr ?

Source: you don 't have to design a whole new set of hospitals to do it .

Reference: Man muss keine neuen Kliniken dafür bauen.

System: Sie müssen eine neue Reihe von Krankenhäusern , es zu tun .

Source: I met a very difficult environment in the pre - elections , an environment which was increasingly polarized , an environment which was shaped by the selfish politics of dominance and exclusion .

Reference: Bei den Vorwahlen traf ich auf ein sehr schwieriges Umfeld, ein Umfeld, das immer stärker polarisierte. Ein Umfeld, das von der egoistischen Politik von Dominanz und Ausschluss geprägt war.

System: Ich traf eine sehr schwierige Umgebung , in der Vor-Wahlen , eine Umgebung

, die zunehmend polarisierten war , war eine Umgebung , die durch die egoistische Politik der Dominanz und Ausgrenzung geprägt .

Source: you see , some people may not do it . they may not get through it .

Reference: Manche werden es vielleicht nicht bis zum Ende schaffen.

System: Sehen Sie , manche Menschen mögen es nicht . Sie können es nicht bekommen .

Source: first question to answer for us : so what ? Reference: Die erste Frage, die ihr uns beantworten müsst, ist: Na und?

System: Erste Frage für uns : Was zu beantworten ?

Source: I will put down the lawsuit and they will send me all the information I ask for .

Reference: Ich ziehe die Anklage zurück, wofür sie mir alle verlangten Informationen schicken.

System: Ich werde die Klage , und sie werden mir all die Informationen , die ich fragen .

Source: so here are some amazing goals that I think we can deal with now .

Reference: Diese erstaunlichen Ziele können wir damit jetzt erreichen:

System: Hier sind einige erstaunliche Ziele , dass ich denke , wir können jetzt behandeln .

Source: I got a chance to come by plane for my first time for TED .

Reference: Ich durfte mit einem Flugzeug anreisen, für meinen ersten TEDTalk.

System: Ich bekam eine Chance von Flugzeug für mein erstes Mal für TED zu kommen .

Source: this is when we started asking passing tourists to take the picture .

Reference: Damals fingen wir an, vorbeilaufende Touristen zu bitten, ein Bild von uns zu machen.

System: Das ist , als wir anfingen , vorbeifahrenden Touristen das Bild zu nehmen .

Source: during Taliban years , I remember there were times I would get so frustrated by our life and always being scared and not seeing a future .

Reference: Ich weiß noch genau, dass ich in den Jahren unter den Taliban manchmal so frustriert war von unserem Leben, von der ständigen Angst und der Perspektivlosigkeit,

System: Während Taliban Jahre , ich erinnere mich , es gab Zeiten , ich würde so frustriert durch unser Leben und immer Angst , und nicht die Zukunft .

Source: you already tried it once , right , and you %-% I tried it before , but I stopped because it gave me a shock .

Reference: Du hast es bereits einmal versucht, nicht wahr – Ja, ich hab's schon mal versucht, aber den Versuch wieder aufgegeben, weil ich ... einen Schlag bekam.

System: Sie versuchte es schon einmal , und ich versuchte es , aber ich hörte auf , weil es gab mir einen Schock .

Source: aware of the devastation and the challenges , I was keen among many other women to rebuild the Libyan civil society , calling for an inclusive and just transition to democracy and national reconciliation .

Reference: Der Verwüstung und den Herausforderungen bewusst, brannte ich, wie so viele andere Frauen darauf, die Zivilgesellschaft Lybiens wieder aufzubauen. Wir forderten einen mitbestimmten und gerechten Übergang zu Demokratie und nationalem Ausgleich.

System: Die Verwüstung und die Herausforderungen , ich war bei vielen anderen Frauen , der libyschen Zivilgesellschaft , für eine integrative und Übergang zur Demokratie und nationale Versöhnung wiederaufzubauen .

Source: also , I realized there was a wide gap between North and South .

Reference: Zudem habe ich den großen Unterschied zwischen Nord- und Südkorea wahrgenommen.

System: Auch merkte ich , dass es eine große Kluft zwischen Nord und Süd .

Source: then you have to offer them absolute , dedicated , passionate service to them .

Reference: Als nächstes müssen Sie ihnen absoluten, engagierten und leidenschaftlichen Service bieten.

System: Dann muss man ihnen absolute , engagierte , leidenschaftliche Dienst .

A.2. Sentences, that have the worstest Bleuscore in the German-EnglishNews translation

Source: um 19 Uhr gestaltet die biraböhmische Blasmusik aus Schömberg den großen Unterhaltung Abend .

Reference: At 19:00, the "biraböhmische" (a play on words alluding to the regions pear trees, Bohemia, and perhaps their love of beer as well!) wind music group from Schömberg, provide a great evening of entertainment.

System: At 7.00 p.m. the biraböhmische brass bands from Schömberg large entertainment evening .

Source: zu sehen sind etwa fünf Meter hohe Ruinen aus Hand Teller großen Steinen .

Reference: Ruins measuring five metres tall, made from stones around the size of the palm of your hand, are waiting to be discovered.

System: To see about five metre high ruins of palm stones .

Source: die Unternehmen Führung prahlt normaler Weise nicht über die Gegen Maßnahmen , die man ergriffen hat , denn das verschafft Angreifern Vorteile .

Reference: Corporate leaders are not generally keen to boast about the countermeasures they have taken, in case this hands an advantage to an attacker.

System: The management does not normally took retaliatory measures , which affords attackers advantages .

Source: sie nannte Sprich Wörter wie : Holz Auge sei wachsam , wie Schuppen von den Augen fallen , ein Auge auf jemand werfen , den Seinen gibt es der Herr im Schlaf , seine Hände in Unschuld waschen .

Reference: She referred to sayings such as: "Holzauge sei wachsam" (keep your eyes peeled), "wie Schuppen von den Augen fallen" (like scales falling from one's eyes), "ein Auge auf jemand werfen" (to cast an eye on someone), "den Seinen gibt's der Herr im Schlaf" (good things come to some when they sleep) and "seine Hände in Unschuld waschen" (to wash one's hands of something).

System: She called sayings like : wooden eye is vigilant , as to the fact that an eye on someone , the gang , there is the Lord in his sleep , wash his hands .

Source: die Kinder Gärten und Kinder Tage Stätten im Zollernalbkreis bekommen in diesen Tagen Post .

Reference: Over the next few days, the nursery schools and day care centres in Zollernalb district will be receiving mail.

System: The kindergartens and nurseries in the Zollernalbkreis post these days .

Source: NSA erklärt , ein " interner Fehler " und nicht Hacker seien Schuld am Crash der Website

Reference: NSA Blames "Internal Error," Not Hackers, For Website Crash

System: NSA declared that an " internal error " and not hackers are to blame for the crash of the website

Source: die neue Hand Gepäck Gebühr von Frontier tritt erst im Sommer in Kraft , ein genaues Datum gibt es noch nicht .

Reference: Frontier's new carry-on fee won't start until summer, though a date hasn't been set.

System: The new cabin baggage fee of Frontier is only in the summer in force , it did not have an exact date .

Source: besonders interessant in Zeiten in denen in Nagold die Geburt Klinik wegrationalisiert werden soll .

Reference: This is particularly interesting at a time when Nagold is set to do away with its maternity clinic.

System: Particularly interesting in times where the birth clinic being in Nagold .

Source: die Verbindlichkeiten gehen dagegen auf die Gesamt Gemeinde über .

Reference: Accounts payable, on the other hand, will be passed over to the "church community" as a whole.

System: The liabilities are opposed to the overall community .

Source: Fahrer , der mit 210 km / h und heißem Getränk zwischen den Beinen gerast war , erhält 1000 £ Strafe

Reference: Driver speeding at 130mph with hot drink between legs fined £1,000

System: Driver , who was with 210 km / h , and hot drinks between the legs gerast gets £ 1000 fine

Source: Australierin legt Berufung gegen Gefängnis Strafe in Thai Land ein

Reference: Australian woman appeals Thai jail time

System: Australierin sets appeal against prison sentence in Thailand

Source: Passagiere können auch ausschließlich beim Direkt Kauf auf der Website von Frontier Plätze reservieren .

Reference: It also allows passengers to choose their seat in advance only if they buy directly from the Frontier website.

System: Passengers can also book exclusively to the direct purchase on the website of Frontier .

Source: der feierliche Ernst , mit dem der Film erzählt wird , macht ihn auch für das Publikum zum quäle rischen Büßerhemd .

Reference: The earnest solemnity of the storytelling risks making it a hair shirt-like ordeal for audiences, too.

System: The solemn seriousness with which the film is told , makes him the nasty hair shirt for the audience .

Source: Schiltach muss dafür 220000 Euro in die Hand nehmen .

Reference: Schiltach will have to contribute up to EUR 220,000 to the project.

System: Schiltach must take 220000 euros .

Source: auch für die Kern Stadt gelte , besonders nicht mobile Mitbürger müssten einen Brief Kasten mit Sonntag Leerung auch fußläufig erreichen können .

Reference: It is also the case for the town centre that those citizens in particular who are not mobile must also be able to access a post box with a Sunday collection on foot.

System: Also applies to the core of the city , especially mobile citizens should have a letterbox with Sunday collection can reach within walking distance .

Source: ich bin mir sicher , dass Fentons süße Kleidung ihren Teil dazu beigetragen hat , das passte wirklich perfekt zum Anlass .

Reference: I'm sure Fenton was helped by his cute clothes, he really looked the

part.

System: I am sure that Fentons sweet clothes has played its part , the truly perfectly matched the occasion .

Source: der Text vermittelt sich auf diese angestrengte Weise jedoch kaum .

Reference: However, the source text makes barely any reference to this intense delivery.

System: The text on the hard way .

Source: die gestiegenen Kosten pro Einheit sind immer noch geringer als bei derzeit verwendeten Maschinen mit zehn Sitzen pro Reihe .

Reference: The gain in unit costs is blunted compared with 10-abreast now in use.

System: The increased costs per unit are still lower than currently used machines with ten seats per row .

Source: der Feuer Wehr gelang es , ein Hündchen in Sicherheit zu bringen , das auf einem 15 Meter hohen Vorsprung in einer Fels Wand festsaß .

Reference: A puppy had a lucky escape after fire crews were called to lift her to safety when she somehow got herself stuck 50ft up on a precarious cliff ledge.

System: Firefighters managed to bring a little dog in security , which was stranded on a 15-meter-high ledge in a rock wall .

Source: es standen fünf Spiele pro Team auf dem Plan , wobei jeder gegen jeden spielte .

Reference: Five matches were scheduled for each team, with every team playing one match against all the others.

System: There were five games per team on the plan , each played against everyone .

A.3. Sentences, that have the worstest Bleuscore in the German-EnglishTedX translation

Source: es erfindet sich seit den Sechzigern aber trotzdem ständig neu und ist deswegen meiner Meinung nach auch als fester kultureller Bestand Teil zu sehen .

Reference: It came about in the sixties but it 's always renewing itself and so I believe it 's a fixed part of culture .

System: It invents since the 1960s but still constantly new and that is why I think is also cultural as an integral part .

Source: so sagen es jeden falls auch die Kalküls .

Reference: That ' s what those who are calculating the costs are saying , at any rate .

System: Anyway , so it also say the calculations .

Source: ja immer in der Forschung , welches Bezug System beschäftigt dich momentan am meisten .

Reference: Yes , always with the object of researching what frame of reference you are dealing with most at any given moment .

System: You always go into the research , which reference system currently employs the most .

Source: und am Liebsten ist , wenn jeder , der Kunde dieser Bank ist , Miteigentümer dieser Bank ist .

Reference: And the best thing would be for every one of the bank 's clients to be one of the owners too .

System: And most loved ones , if any , the customer is this bank is co-owner of this bank .

Source: die schwimmen halt und sollen dann an ihren ja Gelenken dann über Hydraulik den Strom produzieren .

Reference: They swim and are then supposed to produce electricity in their , well , limbs , via hydraulic power .

System: The stop and then swim to their joints then yes on hydraulics produce

the electricity .

Source: also wir mei maulen jetzt nicht weiter rum , sondern wir sagen , was funktioniert am Fern sehen ?

Reference: So let 's no longer grumble about , let 's say , what does work in TV ?

System: So we now mei maulen not further , but we say what works on television ?

Source: eigentlich geht es so darum '

Reference: Actually , it comes down to ,

System: It is actually so hard '

Source: und deswegen haben wir uns überlegt , wir machen jetzt Folgendes .

Reference: And so we thought we 'd do the following . System: And that is why we said , now , we are going to do this .

Source: ich hab mich beschäftigt dann weiter , was gibt es denn sonst noch ?

Reference: Then I moved on and wondered whether there were any other options .

System: I have been looking further , then what else is there ?

Source: was Olympioniken eigentlich neben Olympia noch so machen .

Reference: What do Olympic athletes actually do , besides competing in the Olympic Games .

System: What is actually still so Olympians alongside Olympic .

Source: hätte eigentlich einen eher theoretischen Eingang gewählt .

Reference: Actually I 'd wanted to start off on a more theoretical note .

System: Should have chosen a rather theoretical entrance .

Source: es ist untergegangen , alle sind tot . trotzdem gucken sie sich den an .

Reference: It sank , they all died . And yet they watched it .

System: It is collapsed , are all dead . They still look to each other .

Source: weil wir ja nicht alle in ausreichender Form zur Verfügung haben .

Reference: Because we don 't have enough of all these materials at our disposal .

System: Yes because we do not all have in sufficient form .

Source: der Elektro Smart oder der Diesel Smart , der dann herum steht ist im Prinzip noch .

Reference: The Electrosmart or Dieselsmart car that 's standing around is basically still .

System: The electric Smart or the diesel Smart , who is still around in principle .

Source: ich selber bin mittlerweile durch meine Arbeit zu einer eigenen Marke geworden und hab fest gestellt , dass das Wort Marke von dem Wort markieren rührt .

Reference: My work has turned me into my own brand over time and I 've asserted that the German word for brand , Marke , is based on the German verb markieren , to mark .

System: I myself am now through my work become an own brand and found that the word brand stems from the word mark .

Source: das war , wozu ? das braucht man eigentlich nicht .

Reference: It was like , why ? You don 't really need that .

System: That was what ? That you do not need actually .

Source: das haben wir ganz genauso .

Reference: It occurs in our country , too .

System: That is what we have in exactly the same way .

Source: da muss man wirklich Angst haben , dass man nicht

Reference: One should really be worried there , that they don 't

System: Because you have to really fear is that you do not

Source: und das Mit jeder Fahrt drüber wird das härter .

Reference: Every time a train rides over it , it gets harder .

System: And this With each ride over the going gets tougher .

Source: also , es muss nicht immer selten sein , damit man sich Gedanken machen kann , wie bringt man es raus ?

Reference: So it doesn 't have to also be rare for one to think about how to excavate it ?

System: So , it does not have to be always rare , so you can be concerned , how do you get it out ?

A.4. Sentences, that have the worstest Bleuscore in the German-EnglishTed translation

Source: manche werden es vielleicht nicht bis zum Ende schaffen .

Reference: You see, some people may not do it. They may not get through it.

System: Some are perhaps not going to make it to the end .

Source: ich befürchte , dass die Veränderungen nicht von Dauer sind und sich mit dem Abzug der US - Truppen wieder alles ändert .

Reference: I fear that these changes will not last much beyond the U.S. troops' withdrawal.

System: I am afraid that the changes are not permanent and with the withdrawal of US troops changes everything again .

Source: wenn etwas extrem billig ist , kann man es auch in großem Maß einsetzen .

Reference: When something becomes ultra-low cost, it becomes massively scalable.

System: If something is extremely cheap , you can also use it on a large extent

.

Source: wir müssen aufhören im Namen des Zorns zu handeln und einen Tag der Rache zu fordern .

Reference: We need to stop acting as agents of rage and calling for days of rage.

System: We must stop to act in the name of anger and to demand a day of revenge .

Source: aber das stimmt nicht . es geht zurück auf einen Befürworter der Presse Freiheit .

Reference: It's not. It's down to a campaigner for the freedom of the press.

System: But that is not true . It goes back to an advocate of press freedom .

Source: Strom Ausfälle wurden immer häufiger , sodass sich nachts um mich alles verdunkelte , außer den Lichtern von China auf der anderen Seite des Ufers , an dem wir wohnten .

Reference: Power outages also became more and more frequent, so everything around me was completely dark at night except for the sea of lights in China, just across the river from my home.

System: Blackouts were frequently at night , so that is all about me darkened , except the lights of China on the other side of the shore , where we lived .

Source: und wenn ihr das beschreibt , was ihr macht , dann vermeidet Fach Wörter .

Reference: And when you're describing your science, beware of jargon.

System: And if you describe what you do , then avoids technical words .

Source: und wie Ballet erfordert es enorm viel Übung .

Reference: And like ballet, it takes an extraordinary level of training.

System: And how ballet requires a tremendous amount of exercise .

Source: mit grünen Bohnen im Wert von einem Dollar erhält man Obst und Gemüse im Wert von 75 Dollar .

Reference: When one dollar's worth of green beans will give you 75 dollars' worth of produce.

System: With green beans in the value of a dollar you get fruit and vegetables in the value of 75 dollars .

Source: wir müssen das Bild von Gang Stern umdrehen .

Reference: We gotta flip the script on what a gangsta is.

System: We need to turn around the image of gangsters .

Source: was geschieht , wenn jemand zusticht ? ein anderer wird erstochen , und mit diesen Punkten erstecht ihr euer Publikum .

Reference: What do bullets do? Bullets kill, and they will kill your presentation.

System: What happens when someone zusticht ? Another one gets stabbed , and with these points erstecht her your audience .

Source: sie waren nicht monarchistisch , sie wollten den König für eine bestimmte Zeit Spanne wählen .

Reference: And they weren't even being monarchical there, they had the idea that you could be elected King for a fixed term.

System: They were not monarchistisch , they wanted to choose the king for a certain period of time .

Source: ebenso wenig . die Infrastruktur ist schon da .

Reference: The neurologist's test is non-invasive. They both use existing infrastructure.

System: Just as little . The infrastructure is already there .

Source: seitdem habe ich sieben Häuser in der Nachbarschaft mit den Lichtern versorgt und sie funktionierenen wirklich gut .

Reference: Since now, I've set up seven homes around my community, and they're really working.

System: Since then , I have seven houses in the neighborhood with the lights , and they are working really well .

Source: König nennen , und ihn womöglich noch auf Ideen zu seiner Nachfolge bringen .

Reference: They didn't want to call him King in case that gave him ideas, or his successor ideas.

System: King , and perhaps even to bring ideas to his successor .

Source: deshalb wurde uns gesagt : " meine Gnade hat Vorzug vor meinem Groll . "

Reference: Thus we are told that "My mercy takes precedence over my anger."

System: That is why we were told : " my grace has advantage in front of my grievance . "

Source: Fach Wörter sind ein Hindernis zum Verständnis eurer Gedanken .

Reference: Jargon is a barrier to our understanding of your ideas.

System: Technical words are an obstacle to understanding your thoughts .

Source: nur so konnten wir beide zur Schule gehen .

Reference: It was the only way we both could be educated.

System: Just so we could both go to school .

Source: ich habe sogar meine Freunde mit nach Hause gebracht und gemeinsam installieren wir die Lichter dort , wo es noch keine gibt , und ich zeige den Menschen , wie man sie benutzt .

Reference: I even took my friends back to my community, and we're installing the lights to the homes which don't have [any], and I'm teaching them how to put them.

System: I have even brought my friends home and together we install the lights where there are no , and I am going to show the people how to use them .

Source: als erstes muss ihnen Diskretion angeboten werden .

Reference: First, you have to offer them confidentiality.

System: First of all , they must be offered discretion .