**MACHINE TRANSLATION**

**By**

**SARTHAK GUPTA**

**A Project Report submitted to**

**University at Albany**

**in partial fulfillment of the requirements**

**for the degree of**

**MASTER OF SCIENCE**

**in**

**Computer Science and Engineering**

**------------------------------------**

**Prof.Dr. Tomek Strzalkowski**

tomek@albany.edu

[CSdept@albany.edu](mailto:CSdept@albany.edu)

-------------------------------------------

Date:

**AUTHORIZATION FOR REPRODUCTION**  
**OF PROJECT**

I grant permission for the reproduction of this project in its entirety, without further authorization from me, on the condition that the person or agency requesting reproduction, absorb the cost and provide proper acknowledgment of authorship.

Date SARTHAK GUPTA

943 Western Avenue

Albany, NY, 12203

**ABSTRACT**

Machine Translation is a task in Natural language processing (NLP), where the automatic systems are used to translate the text from one language to another language. In this work we implement Word2Vec algorithm to create Machine Translation system where the automatic translation is used to translate from English to Spanish. Word2Vec algorithm is widely used in Auto Machine Translation recently. It is also used in several NLP models such as text classification, text filtering by tags, and word and phrase translation. In this paper, we will introduce the main structure of Machine Translation and Word2Vec, mathematical model with some formulas, and will study the process of Machine Translation from English to Spanish.

**TABLE OF CONTENTS**

[1 Introduction 6](#_Toc531027152)

[2 Previous Work 8](#_Toc531027153)

[3 Word Embeddings 12](#_Toc531027154)

[4 Gensim 16](#_Toc531027155)

[4.1 Gensim Word2Vec 16](#_Toc531027157)

[4.2 Keyed Vectors 16](#_Toc531027159)

[5 The Datasets 18](#_Toc531027160)

[5.1 United Nations Parallel Corpus 18](#_Toc531027162)

[5.2 Data Tokenization 18](#_Toc531027163)

[5.3 Data Cleaning 18](#_Toc531027164)

[6 Machine Translation using Word2Vec 20](#_Toc531027165)

[7 Results 25](#_Toc531027174)

[7.1 Ferrero Jeremy Cross Language Dataset 25](#_Toc531027175)

[7.2 Ground truth 25](#_Toc531027178)

[7.3 Bleu score 25](#_Toc531027180)

[7.4 Evaluation 26](#_Toc531027181)

[8 t-Test 37](#_Toc531027182)

[9 Analysis 40](#_Toc531027186)

[10 Conclusion 41](#_Toc531027187)

[11 References 43](#_Toc531027188)

[**Appendix A:**  47](#_Toc531027190)

[**Appendix B:** 55](#_Toc531027191)

[**Appendix C:** 64](#_Toc531027192)

[**Appendix D:** 74](#_Toc531027193)

**TABLE OF FIGURES**

Figure 1 14

Figure 2 14

Figure 3 16

Figure 4 17

Table 1 18

Table 2 19

Table 3 27

Figure 5 28

Table 4 29

Figure 6 30

Figure 7 31

Table 5 32

Figure 8 33

Table 6 34

Figure 9 35

Figure 10 36

Table 7,Table 8 38

Table 9,Table 10 39

# Introduction

Over the past decades, the tremendous increase in industrial growth has a huge impact on the global machine translation marker which enables content to be available in all regional languages across the globe. Machine translation is one field of computer science which enables automatic translation of sentences or documents from one language to another. The research of this field has begun to develop rule-based systems using bilingual dictionary and some hand coded rules. In 1980’s, when increased computational power was available, the gradual transition from rule based to statistical trend was observed. The statistical models rely on parallel data to build a statistical model which is more generic and cost effective as compared to rule-based systems. Recently, according to global machine translation market analysis forecast the growth of machine translation by 25% from 2014~2019.

The main aspect of machine translation development is to assess the quality of output, therefore, the main studying in machine translation apart from developing basic systems revolves around quality estimation, domain adaptation, system combination, evaluation and post editing. The main workshops have been conducted in top conferences which emphasize on the broad aspects of machine translation but still this problem is not completely solved. Quality estimation is one challenge in any field of study. Use of automatic systems in the real-world applications has always been a topic of debate in different communities. In the machine translation system, the system predicts the quality of translation from the information available in source sentence and its corresponding translation which then decides on the amount of post-editing needed to use the translation in a real application. Through the past decades multiple approaches for solving complex machine translation problems have been studied with each approach having its pros and cons. The system combination approach combines the positives from the corresponding translation of each system thereby performing better than each of the individual systems. A system trained on one domain fails to meet the desired standards of output when used on another domain due to the difference in syntax and lack of domain vocabulary. A new system can be trained on new domain, but this requires enough data to be available for the new domain. Due to increasing number of domains this task is cost inefficient as preparing new resources for any domain requires large amounts of effort. This issue is resolved by domain adaptation wherein the research community proposes a generalized method for adapting the existing systems to a new domain with limited amount of effort. The evaluation is a key step of any system and the researchers need to come up with the right and standard metrics to compare our system against. A couple of evaluation metrics have been proposed in past decades viz. BLEU [1], METEOR [2], NIST [3] , WER (word error rate) and TER [4]. The BLEU metric is the most widely used but there is another machine translation community which prefer METEOR due to its high correlation with human judgement. In the current era, the researchers have shifted the major focus to inducing semantic properties in machine translation and use of complex neural networks to solve this complex problem. English and Spain are respectively reported to be the largest spoken languages in the world. The fact that one of these languages is known by 10% (approx.) of the world population, makes it an ideal pair for translation studies. In past most of the work has been focused on English to Indian language translation systems and have laid their emphasis on Hindi to English translation systems, which also finds its use in real world problems. The major use can be seen in areas of news articles translations and the translation of documents in judicial systems. The hearings and document work in regional law courts are performed in local language, which creates an issue when the case moves to high court where English is the official language. The time taken by the human translators to translate each document in local language to English postpones the justice and weakens the judicial system. Due to lack of quality English-Spain systems and its use in real world becomes our main purpose in this thesis.

# Previous Work

To develop the automatic machine translation system, authors have developed the several useful methods recently. The one advanced method in machine translation research is statistical machine translation system (SMT). The first SMT systems were based on IBM models that use words as nuclear translation units [5]. Different variations and extensions on the IBM models were defined.Word alignments were improved by introducing maximum entropy models.Translation improvements were obtained by adding linguistic information. In [6] , words were lemmatized. Interpolating lemma and word alignment models further improved the results [7] . The idea of a hierarchical lexicon, where a word is represented at varying levels of inflectional specificity, was explored [8] . Data-driven morphology reduction when translating from a more-inflected language to a less-inflected language, was presented [9] . Words may not be the best atomic units for translation. Phrase-based SMT models use phrases instead of words [10] , whereby a phrase is a contiguous multi-word sequence and has no linguistic motivation. Linguistic knowledge can help to improve phrase-based SMT; it can be utilized during pre-processing and post-processing stages. From a search perspective, it is better to integrate these stages into one model. Factored translation models integrate linguistic annotation as factors into an extension of phrase models [11] . More complex factored models have been explored for English-Czech translation [12].The translation process can be broken down into translation and generation steps. Generating morphology when translating from a less-inflected language into a more-inflected language poses various challenges [13].

Syntax-based models were also tried, as a means of overcome in the limitation of phrase-based SMT [14]. It has been shown that, in some cases, tree-based models can outperform phrase-based models when dealing with languages with complex morphology. However, while a Slovenian parser is now available [15], the Serbian language still lacks the necessary linguistic infrastructure for syntactic parsing; therefore, this approach cannot be performed on the Slovenian Serbian language pair. Machine translation between closely-related languages has been studied for inflectional, analytical and agglutinative languages.Translating a series of letters instead of words improved the results for the Catalan-Spanish 109 language pair [16].

The other machine translation system in advanced natural processing is neural machine translation system (NMT). This MT system is an end-to-end neural network-based approach to statistical machine translation which has recently been able achieve state-of-the-art translation quality on many language pairs.[17]

Statistical machine translation (SMT) systems maximize the conditional probability p(y | x) of a correct target translation y given a source sentence x. This is done by maximizing separately a language model p(y) and the (inverse) translation model p(x | y) component by using Bayes’ rule:

p(y | x) ∝ p(x | y)p(y)

This decomposition into a language model and translation model are meant to make full advantage of available corpora, i.e. monolingual corpora for fitting the language model and parallel corpora for the translation model. SMT systems tend to model log p(y | x) directly by linearly combining multiple features by using a so-called log-linear model:

where fj is the j-th feature based on both or either of the source and target sentences, and C is a normalization constant which is often ignored. These features include, for instance, pair-wise statistics between two sentences/phrases. The log-linear model is fitted to data, in most cases, by maximizing an automatic evaluation metric other than an actual conditional probability, such as BLEU.

Neural machine translation, on the other hand, aims at directly optimizing log p(y | x) jointly with the feature extraction as well as the normalization constant. This is typically done under the encoder-decoder framework consisting of neural networks.The first network encodes the source sentence x into a continuous-space representation from which the decoder produces a target translation sentence. By using RNN architectures equipped to learn long-term dependencies such as gated recurrent units (GRU) or long short-term memory (LSTM),[18,19] the whole system can be trained end-to-end. Once the model learns the conditional distribution or translation model, given a source sentence the machine translation system can find a translation that approximately maximizes the conditional probability using, for instance, a beam search algorithm.

A large part of the recent successes of NMT has been due to the availability of large amounts of high quality, sentence aligned corpora. In the case of low resource language pairs however, or in a task with severe domain restrictions, large parallel corpora may not be available. In contrast, monolingual corpora are almost always abundant and universally available. Despite being “unlabeled”, monolingual corpora still exhibit rich linguistic structure that can be useful for machine translation.

In this paper we consider the wor2vec model which is widely used in natural language processing. In the below we introduce the structure of word2vec model with mathematical formulas.

# Word Embeddings

Word embedding is one of the most popular representation of document vocabulary. It can capture context of a word in a document, semantic and syntactic similarity, relation with other words, etc. the meaning of the word embedding is that they are vector representations of a word. There are two method represent word vector. One is sparse vector, and another is dense vector.

To explain the sparse vectors, we assume we have a set of 1000 documents, consisting of a total of 5000 unique words. In a very native fashion, we can simply count the number of occurrences of each word in every document, and then represent each word by this 1000-dimensional vector of counts. In our case we have taken 100-dimensional vector of counts using Gensim.

Similarly, consider a large corpus of text with 5000 unique words. Now take a window of some fixed size and for each word pair, we count the number of times it occurs in the window. These counts form a term-term matrix, also called a co-occurrence matrix which in this case will be a 5000x5000 matrix (with most cells 0 if the window size is relatively small). The co-occurrence matrix is not the best measure of similarity between 2 words since it is based on the raw frequency, and hence is much skewed. Instead, it would be desirable to have a quantity which measures how much more likely is it for 2 words to occur in a window, compared with pure chance. This is exactly what PMI measures.

If PMI is positive, the (x,y) pair is more likely to occur together than pure chance, and vice versa. However, a negative value is unreliable since it is unlikely to get many co-occurrences of a word pair in a small corpus. To solve this problem, we define a positive PMI (PPMI) as

This is composed of 2 parts: TF, which denotes the count of the word in a document, and IDF, which is a weight component that gives higher weight to words occurring only in a few documents and hence are more representative of the documents they are present in, in contrast to words like ‘the’ which are present in large number of documents.

Here, *N* is the total number of documents and *dfi*is the number of documents in which word *i*occurs.

The problem with sparse vectors is the curse of dimensionality, which makes computation and storage infeasible. For this reason, we prefer dense vectors, with real-valued elements. Dense vector semantics fall into 2 categories: matrix factorization, and neural embedding.

This is basically a dimensionality reduction technique where we find the dimensions with the highest variances. Suppose we have the co-occurrence matrix A of size , then it is possible to factorize A into:

where *r* is the rank of matrix A (i.e. *r* = maximum number of linearly independent vectors that can be used to form A). Also, *U*is a matrix of the eigenvectors of A and S is a diagonal matrix comprising its eigenvalues. If we rearrange the columns in *U*to correspond with a decreasing order of eigenvalues, we can keep the first *k*columns which will represent the dimensions in the latent space which have the highest variance. These will give us a *k*-dimensional representation for each of the *m*words in the vocabulary.

However, this approach is still constrained since the matrix factorization of A, which may be a large matrix, is computationally complex.

To overcome this problem, we can treat each element in the vector as a parameter to be updated while training a neural network model. We start with a randomly initialized vector and update it at each iteration. This update is based on the vectors of the context (window) words. The hypothesis is that such an update would ultimately result in similar words having vectors which are closer to each other in the vector space.

Word2Vec is one of the most popular technique to learn word embedding using shallow neural network. Word2Vec is a method to construct such an embedding. It can be obtained using two methods (both involving neural networks): Skip Gram and Common Bag of Words (CBOW).

CBOW method takes the context of each word as the input and tries to predict the word corresponding to the context.

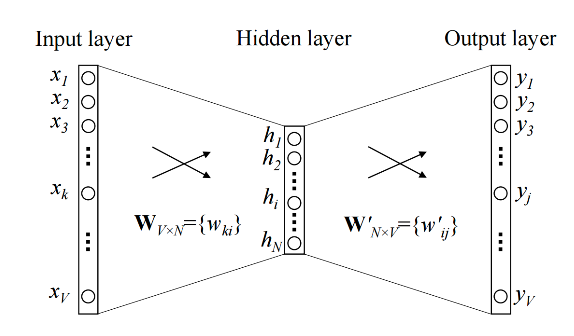


Figure 1 : CBOW Model

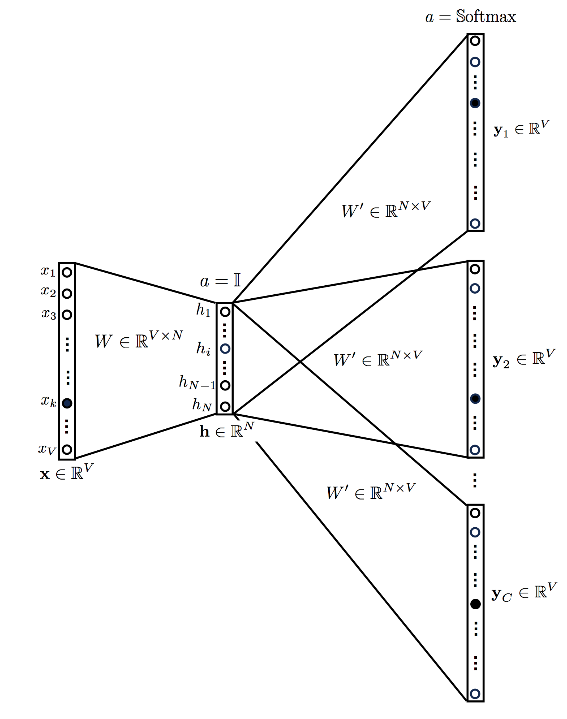


Figure 2 : Skip Gram Model

.

The input or the context word is a one hot encoded vector of size V. The hidden layer contains N neurons and the output is again a V length vector with the elements being the softmax values. The above model takes context words. When *WVN*is used to calculate hidden layer inputs, we take an average over all these context word inputs. So, we have seen how word representations are generated using the context words. But there’s one more way we can do the same. We can use the target word (whose representation we want to generate) to predict the context and in the process, we produce the representations. Another variant, called Skip Gram model does this. In this model we input the target word into the network. The model outputs probability distributions. For each context position, we get probability distributions of V probabilities, one for each word.

Both models have their own advantages and disadvantages. Skip Gram works well with small amount of data and is found to represent rare words well. On the other hand, CBOW is faster and has better representations for more frequent words.

# Gensim

Gensim is a python library that is used to perform word2vec.

## Gensim Word2Vec

The first part is preparing the input for word2vec. The properties of genism are such that it takes only the sequences of sentences as its input with each sentence as a list of words.

Gensim Word2vec accepts several parameters like size, windows, min\_count, workers.

* **Size:** This parameter defines the size of the NN layers or we can say it is the size of the vector representation of each word. The size taken for this project is 100.
* **Windows:** The maximum distance between the current and predicted word within a sentence. I took 5 for my dataset.
* **Min\_count**: The min\_count is to ignore all words with total frequency lower than 1. This means that the word which appears at least once will be considered.
* **Workers:** This is basically for threads to train the model, faster training with multicore machine. The parameters are stored as matrices and the storage capacity is based on the words, their frequency and size

## Keyed Vectors

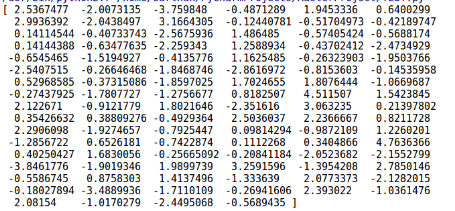


Figure 3: Vector Representation of English Word: ‘Assembly’ of length 100

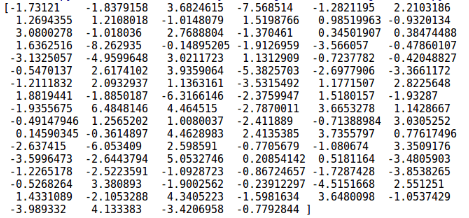


Figure 4: Vector Representation of Spanish Word: ‘Parlamento’ of size 100

# The Datasets

The dataset is collected from the proceedings of the European Parliament from 1996. The corpus consists of about 30 million words for each of the 11 official languages of the Union namely: English, Spanish, Chineses, Danish, Greek, Finnish, French, Italian, Dutch, Portuguese, Swedish. The sources of the parallel text are from multinational institutions and from the various government of the Asian countries.

Only English and Spanish words are taken from the corpus for this project.

## United Nations Parallel Corpus

The dataset was chosen from United Nations proceedings available at Linguistic Data Consortium (LDC), along with which I also selected some talk sentences. All the data in English was line by line converted to Spanish in the training dataset. The purpose of this corpus was to give access of multilingual language documents for various tasks.

## Data Tokenization

The words in English have been tokenized such as words like “can’t” became “can not”, “the man’s” became “the man ‘s”.

## Data Cleaning

 All the words have been converted to lower case so that the lowercase and uppercase of same word are not considered different. The empty lines are removed and their correspondence. Lines which contains XML tags are also removed.

|  |  |  |
| --- | --- | --- |
| **Language** | **Number of Sentences** | **Number of Words** |
| **English** | **1,965,734** | **51,575,748** |
| **Spanish** | **1.965,734** | **49,093,806** |

Table 1: United Nations Proceedings Dataset

Spanish alphabets like "a`, e`, I`, o`, u`" were replaced by English alphabets as these alphabets were not ASCII supported characters.

The data pre-processing involves preparing the data to be used for training the model. Spanish and English text files were extracted and then the keyword 'endsen' was added in every sentence of both the Spanish and English files. This was done to put a stop tag so that the model know that the sentence has ended and is ready for the translation.

For Example: Suppose the sentence in English is: 'My best friend'. Now I will add 'endsen' in the sentence using data\_strip.py and the sentence will be 'My best friend ‘endsen'. The ‘endsen’ in this sentence shows that the 'friend' is the last word in the sentence 'My best friend’.

|  |  |
| --- | --- |
| **No.** | **Types** |
| **1.** | $,%,@,#,+ |
| **2.** | ,.''":=/\; |
| **3.** | [{( )}] |
| **4.** | Single characters- {b}, {c}…. {y}. {z} |
| **5.** | Numeric characters like: 0-9 |
| **6.** | Remove extra white spaces |
| **7.** | Remove trailing and leading spaces |
| **8.** | Remove paragraph spacing |
| **9.** | Remove non-ASCII characters |

Table 2: Characters and Symbols Removed from the text files

# Machine Translation using Word2Vec

In this section, we implement word2vec model. We will consider the word2vec model with input, hidden and output layer and its relationship first.

In the above, we already introduced the word2vec model and its structure. In Figure 1, V is the vocabulary size and the N is the hidden layer size. The input vector is one-hot encoded, that is, that some and all other for . The input vector in our case is the vector representation of English words which is 1× 100.

The weights between the input layer and the hidden layer can be represented by the V ×N matrix W. Each row of W is the N-dimensional vector representation of the associated word in the input layer. Give a context (here a single word), and assuming again and for (one-hot encoding), then

This essentially copies the k-th row of W to h. Here is the vector representation of the input word wI . Note that this implies that the link function of the hidden units is linear (g(x) = x). Recall that

and

So that

From the hidden layer to the output layer there is a different weight matrix which is a N×V matrix. Using this we can compute a score for each word in the vocabulary.

where is the j-th column of the matrix . Note that the score is a measure of the match between the context and the next word and is computed by taking the dot product between the predicted representation () and the representation of the candidate target word ( ).

Now we can use a softmax (log linear) classification model to obtain the posterior distribution of words (turns out to be multinomial distribution).

where yj is the output of the j-th unit of the output layer. Therefore, we get the following formula.

Now, we will consider how to update weights between hidden layer and output layer. This purpose of this step is to maximize , the conditional probability of observing the actual output word wo (denote its index in the output layer as j\*) given the input context word wI (and about the weights). That is,

where is our loss function (which we want to minimize), and j\* is the index of the actual output word (in the output layer).Note (again) that this loss function can be understood as a special case of the cross-entropy measurement between two probabilistic distributions.

The next step is to derive the update equation of the weights between hidden and output layers. Take the derivative of E with respect to j-th unit’s net input, we obtain.

where , the indicator function, i.e., will be 1 when the j-th unit is the actual output word or in our case Spanish word, otherwise . Interestingly this derivative is simply the prediction error of the output layer.

The next step is to take the derivative on to obtain the gradient on the hidden → output weights which (by the chain rule) is

Sine

Now, using gradient descent, we obtain the weight updating equation for hidden → output weights.

and

where η > 0 is the learning rate (standard SGD), , is the i-th unit in the hidden layer, and is the output vector for word .

Next, we will consider backward process. Backward part is to update equations for , we can look at W. Here we take the derivative for E on the output of the hidden layer.

where is the output of the i-th unit in the hidden layer, is defined in Equation 7 (the input of the j-th unit to the output layer), and is the prediction error of the j-th word in the output layer. EH is an N-dimensional vector is the sum of the output vectors of all words in the vocabulary, weighted by their prediction error . The next job is to take the derivative of E with respect to W. So, first recall that the hidden layer performs a linear computation on the values from the input layer, specifically.

Now we can use the chain rule to get the derivative of E with respect to W, as follows

and so

From this we obtain an V × N matrix. Notice that since only one component of x is nonzero, only one row of is non-zero, and the value of that row is EH ( 1×N dimensional vector).

Now we can write the update equation for W as

Here is a row of W (namely the input vector of the (only) context word.

# Results

## Ferrero Jeremy Cross Language Dataset

This dataset is a multilingual in French, English, and Spanish. The datasets contain number of sub-corpora. The corpora chosen for the project is the PAN-PC-11. The PAN-PC-11 corpus contains 26,939 documents and 61,064 plagiarism cases. The focus will be only on parallel aligned English-Spanish translated plagiarism cases. The machine translator is trained on UN dataset and using the PAN-PC-11 dataset.

## Ground truth

The purpose of using the Ground truth is to have a reference for evaluating the translation from the Google translator and Machine translator.

Ground truth is set of 50 random English sentences each from UN Dataset and Ferrero Dataset extracted and their Spanish translation are taken as reference to compare with the same set of English sentences translated using Machine translator and Google translator

## Bleu score

The BLEU score is the metric which is used for evaluating the translations. The Bilingual Evaluation Score, or BLEU for short, is a metric for evaluating a generated sentence to a reference sentence. This metric was firstly proposed by Kishore Papineni et al. in their 2002 paper “[BLEU: a Method for Automatic Evaluation of Machine Translation](http://www.aclweb.org/anthology/P02-1040.pdf)“. This metric works by counting matching n-grams in the candidate translation to n-grams in the reference text, where 1-gram or unigram would be each token and a bigram comparison would be each word pair. The comparison is made regardless of word order.

The central idea behind using BLEU (Bilingual Evaluation Understudy) is to measure how close a machine translation is to a professional human translation. The BLEU metric measures translation on a scale of 0 to 1. The closer is the value to 1, the more the translation correlates to a human translation.

## Evaluation

We use BLEU method to evaluate our machine translation. 50 sentences from Ferrero dataset and 50 sentences from UN dataset is taken as an input/source for the machine translation. These sentences are first translated into output language which is in our case Spanish by our machine translation system and same sentences is translated through Google translator.

The system is run through many epochs and the BLEU score is calculated for every epoch. In this project, we have taken two cases. The first case is when the machine trains for 10 epochs and the second case is when the machine is trained for 20 epochs.

The BLEU score for all the 100 sentences are calculated. After this the BLEU score are compared by mapping them with number of words in the sentence. The reason we are mapping them based on number of words is because we want to check how effective the machine in translating sentences of different length is.

The process of evaluation involves comparing the Google translated sentences with the Ground Truth as a reference and comparing the Machine Translated sentences with the Ground Truth. The difference in comparison of these BLEU score will give us the significance of our system with the Google translator.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SOURCE**  **(Ferrero Dataset)** | **CANDIDATE**  **(MT Epoch 10)** | **BLEU**  **(MT Epoch 10)** | **CANDIDATE**  **(Google Translator)** | **BLEU**  **(Google Translator)** | **REFERENCE**  **(GROUND TRUTH)** |
| Jaime was not behind this | no ha sido fácil ella | 0.40 | Jaime no estaba detrás de esto | 0.60 | Jaime no fue tras ella |
| he remained motionless in the loneliness of the pine insensitive to everything around you like a legendary hero under a spell | ¡viva la historia de darfur | 0.20 | permaneció inmóvil en la soledad del pino insensible a todo lo que te rodeaba como un héroe legendario bajo un hechizo | 0.85 | Permaneció inmóvil en la soledad del pinar, insensible a cuanto le  rodeaba, como un héroe de leyenda sometido a un encantamiento |
| then he passed his hand over his face as if he were waking up by coordinating his ideas | luego esto por el<UNK> <UNK> <UNK> <UNK> a su pregunta | 0.24 | luego se pasó la mano por la cara como si despertara coordinando sus ideas | 0.85 | Luego se pasó una mano por el rostro, cual si despertase, coordinando sus ideas |
| it was the result of his trip to the city back to the civilized life that had upset his solitary and quiet passions of the past the conversation of the young soldiers who lived with the thought put in the woman | ha pasado el siglo pasado historia de iraq y <UNK> <UNK> <UNK> homenaje a la gente | 0.11 | fue el resultado de su viaje a la ciudad de regreso a la vida civilizada que había trastornado sus pasiones solitarias y tranquilas del pasado la conversación de los jóvenes soldados que vivían con el pensamiento puesto en la mujer | 0.87 | Era el resultado de su viaje a la ciudad, la vuelta a la vida civilizada que había trastornado su calma de solitario, despertando pasiones de antaño; la conversación de los jóvenes militares, que vivían con el pensamiento puesto en la mujer |

*Table 3: Google Translation (translation of the English sentence in the Ferrero dataset) and the machine translation with epoch 10 of the same sentences vs Ground Truth (see Appendix A for more examples)*

Figure 5: BLEU Score comparison using Google Translator v/s MT with Epoch 10

The above chart shows us the BLEU score calculated taking 50 sentences from Ferrero Dataset. The reference taken when calculating BLEU score are the Ground Truth translation of the same input sentences. The chart represents the change in the value of BLEU score based on the length of sentences taken as input to our machine. The two samples of points are the two set of sentences translated by the machine translator and the Google translator. The blue points show the BLEU score of the sentences translated by Google Translator and the red point shows the BLEU score of the sentences translated by Machine translation trained for 10 epochs. We can observe from the chart that most of the red points (BLEU score) lies in the range 10% to 40% which is very low as compared to the blue points which lies in the range 60% to 100%. The machine performed better when sentences less than 10 words were given as input. The BLEU for them went upto 60 %. The sentences with words more than 30 have very low BLEU score stating that the machine does not translate long sentences accurately. The Google translator on the other hand has translated better than the machine where the BLEU score is on the higher side for the same input sentences. The Google translator translates the short and long sentences far better than the machine translation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SOURCE**  **(Ferrero Dataset)** | **CANDIDATE**  **(MT Epoch 20)** | **BLEU**  **(MT Epoch 20)** | **CANDIDATE**  **(Google Translator)** | **BLEU**  **(GoogleTranslator)** | **REFERENCE**  **(GROUND TRUTH)** |
| Jaime was not behind this | no se ha referido a tras ella | 0.60 | Jaime no estaba detrás de esto | 0.60 | Jaime no fue tras ella |
| he remained motionless in the loneliness of the pine insensitive to everything around you like a legendary hero under a spell | <UNK> de la | 0.09 | permaneció inmóvil en la soledad del pino insensible a todo lo que te rodeaba como un héroe legendario bajo un hechizo | 0.85 | Permaneció inmóvil en la soledad del pinar, insensible a cuanto le  rodeaba, como un héroe de leyenda sometido a un encantamiento |
| then he passed his hand over his face as if he were waking up by coordinating his ideas | luego de hecho se trata de una forma de que se <UNK> a la comisión sus armas | 0.42 | luego se pasó la mano por la cara como si despertara coordinando sus ideas | 0.85 | Luego se pasó una mano por el rostro, cual si despertase, coordinando sus ideas |
| it was the result of his trip to the city back to the civilized life that had upset his solitary and quiet passions of the past the conversation of the young soldiers who lived with the thought put in the woman | fue el momento de que la historia de la ciudad húngara <UNK> <UNK> | 0.14 | fue el resultado de su viaje a la ciudad de regreso a la vida civilizada que había trastornado sus pasiones solitarias y tranquilas del pasado la conversación de los jóvenes soldados que vivían con el pensamiento puesto en la mujer | 0.87 | Era el resultado de su viaje a la ciudad, la vuelta a la vida civilizada que había trastornado su calma de solitario, despertando pasiones de antaño; la conversación de los jóvenes militares, que vivían con el pensamiento puesto en la mujer |

*Table 4: Google Translation (translation of the English sentence in the Ferrero dataset) and the machine translation with epoch 20 of the same sentences vs Ground Truth (see Appendix B for more examples)*

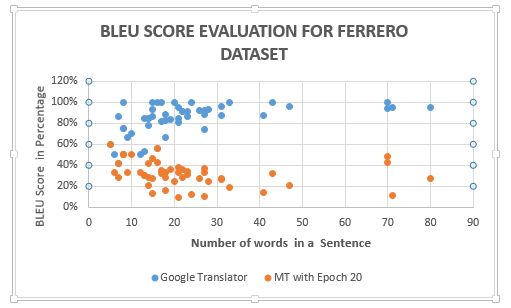


Figure 6: BLEU Score comparison using Google Translator v/s MT with Epoch 20

The above chart shows us the BLEU score calculated taking 50 sentences from Ferrero Dataset. The reference taken when calculating BLEU score are the Ground Truth translation of the same input sentences. The chart represents the change in the value of BLEU score based on the length of sentences taken as input to our machine. The two samples of points are the two set of sentences translated by the machine translator and the Google translator. The blue points show the BLEU score of the sentences translated by Google Translator and the red point shows the BLEU score of the sentences translated by Machine translation trained for 20 epochs. We can observe from the chart that most of the red points (BLEU score) lies in the range 10% to 40% which is very low as compared to the blue points which lies in the range 60% to 100%. The machine performed better when sentences less than 10 words were given as input. The BLEU for them went upto 60 %. The sentences with words more than 30 have very low BLEU score stating that the machine does not translate long sentences accurately. The Google translator on the other hand has translated better than the machine where the BLEU score is on the higher side for the same input sentences. The Google translator translates the short and long sentences far better than the machine translation.

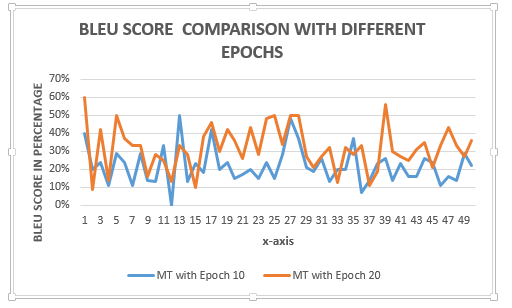


Figure 7: BLEU Score Comparison (Ferrero Dataset) for MT with Epoch 10 vs Epoch 20

Figure 7 shows us how the machine translation accuracy has increased in every sentence. The red line indicates the value of BLEU score for Ferrero Dataset after training the machine for Epoch 20 and the blue line indicates the value after training the machine for Epoch 10.

We can say by looking at the chart that with increase in the Epoch we can increase the accuracy of the machine. The machine’s accuracy can also increase if we train the model with larger dataset.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SOURCE**  **(UN Dataset)** | **CANDIDATE**  **(MT Epoch 10)** | **BLEU**  **(MT Epoch 10)** | **CANDIDATE**  **(Google Translator)** | **BLEU**  **(GoogleTranslator)** | **REFERENCE**  **(GROUND TRUTH)** |
| The House rose and observed a minute silence | la asamblea y el martes <UNK> el martes | 0.33 | La casa se levantó y observó un minute y de silencio | 0.44 | El Parlamento de pie guarda un minuto de silencio |
| I declare resumed the session of the European Parliament adjourned on Friday 17 December 1999 and I would like once again to wish you a happy new year in the hope that you enjoyed a pleasant festive period | visité el parlamento europeo de junio de junio de junio y <UNK> i i böhm i y <UNK> | 0.20 | Declaro reanudado el período de sesiones del Parlamento Europeo aplazado el viernes 17 de diciembre de 1999 y una vez más deseo desearles un feliz año nuevo con la esperanza de que hayan disfrutado de un agradable período festivo | 0.96 | Declaro reanudado el periodo de sesiones del Parlamento Europeo interrumpido el viernes 17 de diciembre pasado y reitero a Sus Senorias mi deseo de que hayan tenido unas buenas vacaciones |
| I believe that the principle of relative stability is a fundamental legal principle of the common fisheries policy and a proposal to subvert it would be legally inadmissible | creo que el documento político de la política de competencia desleal y subsidiariedad debería ser <UNK> | 0.31 | Creo que el principio de estabilidad relativa es un principio jurídico fundamental de la política pesquera común y una propuesta para subvertirlo sería legalmente inadmisible | 0.93 | Creo que el principio de estabilidad relativa es un principio legal fundamental de las politicas pesqueras comunitarias por lo que una propuesta que lo subvierta es legalmente inadmisible |
| This is all in accordance with the principles that we have always upheld | esto se ha hecho de todos los criterios que hemos decidido | 0.36 | Todo esto está de acuerdo con los principios que siempre hemos sostenido | 0.63 | Todo ello de conformidad con los principios que siempre hemos apoyado |
| Now however he is to go before the courts once more because the public prosecutor is appealing | ahora <UNK> la comisión <UNK> <UNK> <UNK> <UNK> | 0.06 | Ahora, sin embargo, debe comparecer ante los tribunales una vez más porque el fiscal está apelando | 0.75 | Sin embargo sucede que va a ser acusado de nuevo pues el fiscal interpondra un recurso |

*Table 5: Google Translation (translation of the English sentence in the UN dataset) and the machine translation with epoch 10 of the same sentences vs Ground Truth (see Appendix C for more examples)*

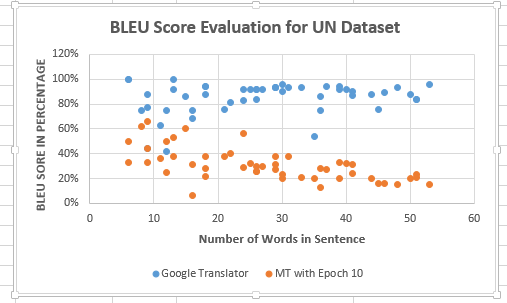


Figure 8: BLEU Score comparison using Google Translator v/s MT with Epoch 10

The above chart shows us the BLEU score calculated taking 50 sentences from UN Dataset. The reference taken when calculating BLEU score are the Ground Truth translation of the same input sentences. The chart represents the change in the value of BLEU score based on the length of sentences taken as input to our machine. The two samples of points are the two set of sentences translated by the machine translator and the Google translator. The blue points show the BLEU score of the sentences translated by Google Translator and the red point shows the BLEU score of the sentences translated by Machine translation trained for 10 epochs. We can observe from the chart that most of the red points (BLEU score) lies in the range 20% to 60% which is very low as compared to the blue points which lies in the range 60% to 100%. The machine performed better when sentences less than 25 words were given as input. The BLEU for them went upto 60 %. The sentences with words more than 30 have very low BLEU score stating that the machine does not translate long sentences accurately. The Google translator on the other hand has translated better than the machine where the BLEU score is on the higher side for the same input sentences. The Google translator translates the short and long sentences far better than the machine translation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SOURCE**  **(UN Dataset)** | **CANDIDATE**  **(MT Epoch 20)** | **BLEU**  **(MT Epoch 20)** | **CANDIDATE**  **(Google Translator)** | **BLEU**  **(GoogleTranslator)** | **REFERENCE**  **(GROUND TRUTH)** |
| The House rose and observed a minute silence | el la cámara se ha reunido y minute silencio<UNK> y <UNK> | 0.44 | La casa se levantó y observó un minute y de silencio | 0.44 | El Parlamento de pie guarda un minuto de silencio |
| I declare resumed the session of the European Parliament adjourned on Friday 17 December 1999 and I would like once again to wish you a happy new year in the hope that you enjoyed a pleasant festive period | declaro reanudado al consejo de abril de abril de 2000 y a la presidencia neerlandesa que se <UNK> y que se aplace un nuevo período de sesiones de 2013 para que podamos llegar a un período de festivo sesiones de 2012 | 0.78 | Declaro reanudado el período de sesiones del Parlamento Europeo aplazado el viernes 17 de diciembre de 1999 y una vez más deseo desearles un feliz año nuevo con la esperanza de que hayan disfrutado de un agradable período festivo | 0.96 | Declaro reanudado el periodo de sesiones del Parlamento Europeo interrumpido el viernes 17 de diciembre pasado y reitero a Sus Senorias mi deseo de que hayan tenido unas buenas vacaciones |
| I believe that the principle of relative stability is a fundamental legal principle of the common fisheries policy and a proposal to subvert it would be legally inadmissible | creo que el principio de política de principio de la política de protección medioambiental de la salud pública es un texto de la que se basa en la propuesta de la comisión | 0.58 | Creo que el principio de estabilidad relativa es un principio jurídico fundamental de la política pesquera común y una propuesta para subvertirlo sería legalmente inadmisible | 0.93 | Creo que el principio de estabilidad relativa es un principio legal fundamental de las politicas pesqueras comunitarias por lo que una propuesta que lo subvierta es legalmente inadmisible |
| This is all in accordance with the principles that we have always upheld | todo esto se ha respetado en todos los principios que se hemos sostenido | 0.78 | Todo esto está de acuerdo con los principios que siempre hemos sostenido | 0.63 | Todo ello de conformidad con los principios que siempre hemos apoyado |
| Now however he is to go before the courts once more because the public prosecutor | sin embargo ahora me parece que se trata de una el forma de censura esta de la ue y de la carta | 0.63 | Ahora, sin embargo, debe comparecer ante los tribunales una vez más porque el fiscal está apelando | 0.75 | Sin embargo sucede que va a ser acusado de nuevo pues el fiscal interpondra un recurso |

*Table 6: Google Translation (translation of the English sentence in the UN dataset) and the Machine translation with epoch 20 of the same sentences vs Ground Truth (see Appendix D for more examples*

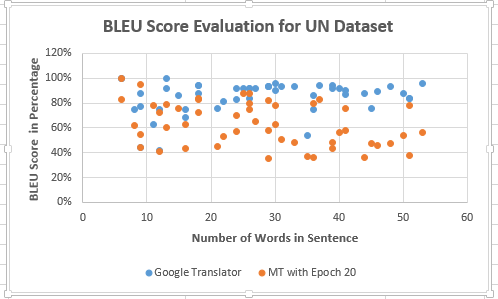


Figure 9: BLEU Score comparison using Google Translator v/s MT with Epoch 20

The above chart shows us the BLEU score calculated taking 50 sentences from UN Dataset. The reference taken when calculating BLEU score are the Ground Truth translation of the same input sentences. The chart represents the change in the value of BLEU score based on the length of sentences taken as input to our machine. The two samples of points are the two set of sentences translated by the machine translator and the Google translator. The blue points show the BLEU score of the sentences translated by Google Translator and the red point shows the BLEU score of the sentences translated by Machine translation trained for 20 epochs.We can observe from the chart that most of the red points (BLEU score) lies in the range 40% to 100% which is very low as compared to the blue points which lies in the range 60% to 100%. We can see from the chart that the BLEU scores for the MT is very close to the Google Translator. The machine performed better when sentences less than 25 words were given as input. The BLEU for them went upto 100 %. The sentences with words more than 30 have low BLEU score stating that the machine does not translate long sentences accurately, but they partially translated it i. e having the BLEU score of 40%. The Google translator on the other hand has translated better than the machine where the BLEU score is on the higher side for the same input sentences. The Google translator translates the short and long sentences far better than the machine translation.

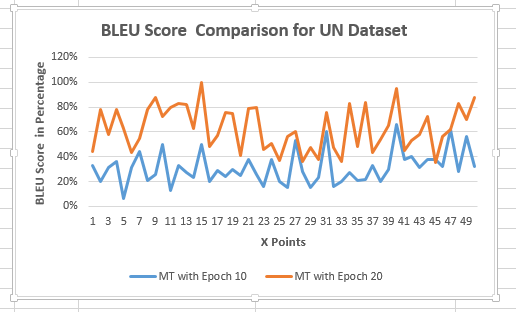


Figure 10: BLEU Score Comparison for MT with Epoch 10 vs Epoch 20

The above chart shows us how the machine translation accuracy has increased in every sentence. The red line indicates the value of BLEU score for United Nation Dataset after training the machine for Epoch 20 and the blue line indicates the value after training the machine for Epoch 10. We can see that the differences increase for the long sentences stating the machine translated better for long sentences than short sentences.

The above shows us the improvement in the translation after increasing the number of Epochs. The accuracy of translation of every sentence has increased dramatically.

# t-Test

The dependent samples being the sentence which are ground truth and the sentences translated by Google Translator and also the sentences translated by the machine.

We will compare the means of two different samples: one being Google Translator and the other being Machine translation. The two sample are generated by calculating the BLEU score taking Ground Truth as a reference. The Ground Truth being the sentences already translated and downloaded.

To perform paired t test we have to know what the null hypothesis is. The null hypothesis is always the accepted facts.

So, the null hypothesis in our case would be

**Null hypothesis (Ho):** Assume that there is no difference and our system does the same translation as Google Translate. In our case value of Ho would be 1. (**Ho = 1)**

**Two Tailed research hypothesis (H1):** There is a significant difference between Google translate and our Machine Translation**. (H1 < > 1).**

The BLEU score for both the Machine translator and Google Translator are calculated taking ground truth as a reference. The average difference is calculated by subtracting the mean BLEU score of Machine translation and Google Translation for different datasets.

The values are taken in percentage to make the calculations easier.

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset Name** | **Epoch Count** | **Average Difference BLEU Score (in percent)** | **Standard Deviation(std)** |
| **Ferrero Dataset** | 10 | 62 | 14.0 |
| **Ferrero Dataset** | 20 | 52 | 12.0 |
| **United Nation Dataset** | 10 | 54 | 19.0 |
| **United Nation Dataset** | 20 | 22 | 15.0 |

Table 7: Average mean Difference and Standard Deviation between Machine Translation and Google Translation for different datasets

Standard Error of the Mean difference:

|  |  |  |
| --- | --- | --- |
| **Dataset Name** | **Epoch Count** | **Standard Error of the Mean difference** |
| **Ferrero Dataset** | 10 | 1.9 |
| **Ferrero Dataset** | 20 | 1.6 |
| **United Nation Dataset** | 10 | 2.6 |
| **United Nation Dataset** | 20 | 2.1 |

Table 8: Standard Error of the Mean Difference between Machine Translation and Google Translation for different datasets

|  |  |  |
| --- | --- | --- |
| **Dataset Name** | **Epoch Count** | **t-score on 49 degrees of freedom** |
| **Ferrero Dataset** | 10 | 0.32 |
| **Ferrero Dataset** | 20 | 0.33 |
| **United Nation Dataset** | 10 | 0.30 |
| **United Nation Dataset** | 20 | 0.10 |

**Degrees of freedom of an estimate is the number of independent pieces of information that went into calculating the estimate. The degree of freedom is (Total number of sentences -1). We are taking 50 sentences and hence the degree of freedom is: 50-1 = 49.**

Table 9: Calculated t-Score for different datasets

|  |  |  |
| --- | --- | --- |
| **Dataset Name** | **Epoch Count** | **p- Value** |
| **Ferrero Dataset** | 10 | 0.75 |
| **Ferrero Dataset** | 20 | 0.74 |
| **United Nation Dataset** | 10 | 0.76 |
| **United Nation Dataset** | 20 | 0.92 |

Table 10: p-Value for different datasets

49 degrees of freedom at an alpha level of 0.05 = 2.01.

Therefore p < 0.05. As the p-value is greater than the alpha level, we can conclude that there is a difference between means. We will reject the null hypothesis (Ho) which states that our system and Google Translator has no significant difference.

**Hence there is a significant difference between Machine Translator and Google Translator and Google Translator is significantly better than Machine Translator.**

# Analysis

In the evaluation of machine translation, we get <UNK> in some translation. The reason of this happen is that the machine translation model allows to produce the <UNK> when it is not sure of the specific target word. This problem is not the problem occurring because the training data is small. Though we trained machine translation model with the large amount of words, we can face this problem. The technique to overcome this problem is widely used in the world. In the papers [41][42], the neural machine translation models were proposed to overcome this problem.

We can see the result of the machine translation model trained with UN data set than the result of machine translation trained with Ferrero data set. For the Ferrero data set, the training result for epoch 20 is better than the result for epoch 10. It means the training is not convergence completely for small epoch. The similar happening is occurring for UN data set. If we train machine translation model with large epoch, perhaps we can get better result. And for same epoch, the result seems better for UN data set than for Ferrero data set. It means the training process with UN data set is well adapted in the machine translation model.

# Conclusion

In this paper, we studied the machine translation model based on word2vec. Through this project, we considered the word2vec model and analysis the structure of those models. Based on the comparison of machine translation models, we designed our machine translation model to translate from English to Spanish and evaluated it. Two data set (UN and Ferrero) are used for training of model in this project. The work for machine translation is increasing with computer since and deep learning algorithms, recently. In the future, we can study our model with more different data set. We also try to analysis the result from several machine translation to get better machine translation model from English to Spanish.

# 11 References

[1] Kishore Papineni, Salim Roukos, Todd Ward, John Henderson, and Florence Reeder. 2002. Corpus-based comprehensive and diagnostic MT evaluation: Initial Arabic, Chinese, French, and Spanish results. In Proceedings of Human Language Technology 2002, San Diego, CA.

[2] Satanjeev Banerjee and Alon Lavie. 2005. METEOR:An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments. In Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization

[3] Doddington, George. (2002). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics.

[4] Matthew Snover, Bonnie Dorr, Richard Schwartz, Linnea Micciulla, and John Makhoul. 2006. A Study of Translation Edit Rate with Targeted Human Annotation. In Proceedings of Association for Machine Translation in the Americas.

[5] Brown, Peter E; Cocke, John; Della Pietra, Stephen A.; Della Pietra, Vincent J.; Jelinek, Frederick; Lafferty, John D.; Mercer, Robert L.; and Roossin, Paul S. (1990). "A statistical approach to machine translation."

[6] Popovic, Maja & Ney, Hermann. (2004). Towards the use of word stems and suffixes for statistical machine translation.

[7] R. Zhang and E. Sumita, “Boosting statistical machine translation by lemmatization and linear interpolation,” in Proc. of the ACL: companion volume, 2007

[8] Niessen, S., Ney, H.: Statistical machine translation with scarce resources using morpho-syntatic information. Computational Linguistics 30, 181–204 (2004)

[9] Brest, Janez & Zamuda, Aleš & Fister, Iztok & Maučec, Mirjam. (2010). Sec. Large scale global optimization using self-adaptive differential evolution algorithm. IEEE World Congress on Computational Intelligence 2010.

[10] Franz Josef Och , Hermann Ney, The Alignment Template Approach to Statistical Machine Translation, Computational Linguistics, December 2004

[11] Koehn, Philipp, and Hieu Hoang. "Factored Translation Models." EMNLP, 2007.

[12] Ales Tamchyna, Petra Galu ˇ sˇcˇakov ´ a, Amir Kamran, Milo ´ sˇ Stanojevic, and Ond ´ ˇrej Bojar. 2012. Selecting data for English-to-Czech machine translation. In Proceedings of the Seventh Workshop on Statistical Machine Translation, Montreal, Canada, June. Association for Computational Linguistics

[13] Littrell, Romie. (2008). Book Review: Minkov, M. (2007). What makes us different and similar: A new interpretation of the World Values Survey and other cross-cultural data. Sofia, Bulgaria: Klasika y Stil Publishing House. ISBN: 978-954-327-023-1, 240 pp. Journal of Cross-Cultural Psychology.

[14] Chris Quirk and Simon Corston-Oliver. 2006. The Impact of Parse Quality on Syntactically-Informed Statistical Machine Translation. In Proceedings of EMNLP

[15] Dobrovoljc, Katarina & Falnoga, Ingrid & Tušek-Žnidarič, Magda & Mazej, Darja & Ščančar, Janez & Bulog, Boris. (2012). Cd, Cu, Zn, Se, and Metallothioneins in Two Amphibians, Necturus maculosus (Amphibia, Caudata) and Bufo bufo (Amphibia, Anura). Biological trace element research.

[16] Vilar, Miguel & Chan, Chim & Santos, Dana & Lynch, Daniel & Spathis, Rita & M Garruto, Ralph & Lum, J. (2013). Vilar et al 2012 Chamorro Origins [1].

[17] C. Gulcehre, O. Firat, K. Xu, K. Cho, L. Barrault, H.-C. Lin, F. Bougares, H. Schwenk, Y. Bengio, On using monolingual corpora in neural machine translation, arXiv preprint arXiv:1503.03535.

[18] I. Sutskever, O. Vinyals, Q. Le, Sequence to sequence learning with neural networks, in: Advances in Neural Information Processing Systems (NIPS 2014), 2014.

[19] K. Cho, B. van Merrienboer, C. Gulcehre, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, in: Proceedings of the Empiricial Methods in Natural Language Processing (EMNLP 2014), 2014, to appear.

[20] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, arXiv preprint arXiv:1409.0473.

[21] S. Jean, K. Cho, R. Memisevic, Y. Bengio, On using very large target vocabulary for neural machine translation, arXiv preprint arXiv:1412.2007.

[22] Allauzen, Alexandre, et al. “LIMSI@ WMT11.” Proceedings of the Sixth Workshop on Statistical Machine Translation. Association for Computational Linguistics, 2011.

[23] Auli, Michael, et al. “Joint Language and Translation Modeling with Recurrent Neural Networks.” EMNLP. 2013.

[24] Bengio, Yoshua, Aaron Courville, and Pascal Vincent. “Representation learning: A review and new perspectives.” Pattern Analysis and Machine Intelligence, IEEE Transactions on 35.8 (2013): 1798-1828.

[25] Le, Hai-Son, et al. “LIMSI@ WMT’12.” Proceedings of the Seventh Workshop on Statistical Machine Translation. Association for Computational Linguistics, 2012.

[26] Mikolov, Tomas, et al. “Recurrent neural network-based language model.”INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September 26-30, 2010. 2010.

[27] Schwenk, Holger, Anthony Rousseau, and Mohammed Attik. “Large, pruned or continuous space language models on a gpu for statistical machine translation.”Proceedings of the NAACL-HLT 2012 Workshop: Will We Ever Really Replace the N-gram Model? On the Future of Language Modeling for HLT. Association for Computational Linguistics, 2012.

[28] S´ebastien Jean, Kyunghyun Cho, Roland Memisevic, and Yoshua Bengio. 2015. On using very large target vocabulary for neural machine translation.

[29] N. Kalchbrenner and P. Blunsom. 2013. Recurrent continuous translation models. In EMNLP.

[30] Jun Suzuki and Masaaki Nagata. 2017. Rnnbased encoder-decoder approach with word frequency estimation. CoRR abs/1701.00138. <http://arxiv.org/abs/1701.00138>.

[31] Sébastien Jean, Kyunghyun Cho, Roland Memisevic, and Yoshua Bengio. 2015. On using very large target vocabulary for neural machine translation. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, pages 1–10. <https://doi.org/10.3115/v1/P15-1001>.

[32] Ke Tran, Arianna Bisazza, and Christof Monz. 2014. Word translation prediction for morphologically rich languages with bilingual neural networks. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, pages 1676–1688. <https://doi.org/10.3115/v1/D14-1175>.

[33] Fandong Meng, Zhengdong Lu, Hang Li, and Qun Liu. 2016. Interactive attention for neural machine translation. CoRR abs/1610.05011. <http://arxiv.org/abs/1610.05011>.

[34] Haitao Mi, Zhiguo Wang, and Abe Ittycheriah. 2016. Vocabulary manipulation for neural machine translation. CoRR abs/1605.03209. <http://arxiv.org/abs/1605.03209>.

[35] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A simple way to prevent neural networks from overfitting.

[36] Hardmeier, C., Volk, M., 2009. Using linguistic annotations in statistical machine translation of film subtitles, in: Proceedings of the the 17th Nordic Conference of Computational Linguistics, Netherlands, pp. 57–64.

[37] Homola, X., Kuboˇn, V., Viˇciˇc, J., 2009. Shallow transfer between Slavic language, in: Recent Advances in Intelligent Information Systems, pp. 219–232.

[38] Ittycheriah, A., Roukos, S., 2005. A maximum entropy word aligner for Arabic-English machine translation, in: Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Vancouver, Canada, pp. 89–96.

[39] Kirik, H., Fishel, M., 2008. Modelling linguistic phenomena with unsupervised morphology for improving statistical machine translation, in: Proceedings of the SLTC08 Workshop on Unsupervised Methods in NLP, Stockholm, Sweden.

[40] Koehn, P., 2010. Statistical Machine Translation. Cambridge University Press, Cambridge, UK.

[41] <https://www.aclweb.org/anthology/D13-1176>

[42] <https://arxiv.org/abs/1409.3215>

**Appendix A**

**Sentence Translation from Machine Translator with Epoch 10 and Google Translator for Ferrero Dataset**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| What mattered was that he knew what so many times he had thought vaguely in the isolation of the tower without being able to fulfill his wishes | qué duda lo sabía de que nunca ha cambiado de su situación que quería escucharla | 0.24 | Lo que importaba era que supiera lo que tantas veces había pensado vagamente en el aislamiento de la torre sin poder cumplir con sus deseos | 0.88 | Lo importante era que conociese lo que tantas veces había pensado él vagamente en el aislamiento de la torre, sin poder dar forma precisa a sus deseos |
| He continued his way slowly without reaching the family | usted puede permitir que la ue no puede <UNK> | 0.11 | siguió su camino lentamente sin llegar a la familia | 0.50 | Continuó lentamente su camino, para no alcanzar a la familia de Can |
| he joined his mother and brother | ha dado su trabajo y la comisión | 0.29 | se unió a su madre y hermano | 0.66 | se había reunido con su madre y su hermano |
| observed from a height when the group now walked towards the valley towards the farm | murió de la historia de <UNK> <UNK> <UNK> en contra de todos | 0.14 | observó desde una altura cuando el grupo caminó ahora hacia el valle hacia la granja | 0.66 | Los vio desde una altura, cuando el grupo caminaba ya por el valle con dirección a la alquería |
| twisted went to the pirate's tower, but when I got closer, he continued his way without stopping until the marcamino avoiding the approach | la espera de su libro puede ser la excelencia | 0.13 | retorció fue a la torre del pirata, pero cuando me acerqué, siguió su camino sin detenerse hasta el marcamino evitando el enfoque | 0.42 | torció su marcha, evitando aproximarse a Can |
| He went to the pirate's tower, but when I got close, he continued his way without stopping to the sea | fue lo sabía el pacto pero <UNK> <UNK> el reino unido | 0.33 | fue a la torre del pirata, pero cuando me acerqué, siguió su camino sin detenerse hasta el mar | 1.00 | Fue hacia la torre del Pirata, pero al llegar cerca de ella continuó su camino, no deteniéndose hasta el mar |
| this secular work was nibbling the coast snatching its stone shell sheet by leaf | sk señorías es muy importante | 0.00 | este trabajo secular estaba mordisqueando la costa arrebatando su hoja de concha de piedra por hoja | 0.86 | Esta labor secular iba royendo la costa, arrebatándola su coraza de piedra, lámina por lámina |
| fragments as walls | de ella señora <UNK> <UNK> <UNK> <UNK> <UNK> | 0.30 | fragmentos lo como muros | 0.50 | de ella fragmentos enormes como murallas |
| first forming an imperceptible slit that got bigger with the course always | ¡viva la historia de una nación de francia <UNK> <UNK> <UNK> | 0.13 | se primero formando una hendidur imperceptible que se hizo más grande con el curso siempre | 0.85 | primeramente formando una rendija imperceptible, que se agrandaba con el curso de los siglos |
| he was to sit on the edge of a large rock he advanced a fragment of rock separated from | se ha hablado de un mundo internacional como un ejemplo de <UNK> | 0.23 | debía sentarse en el borde de una gran roca avanzó un fragmento de roca separada de | 0.74 | fue a sentarse en el borde de un gran peñasco avanzado, de un fragmento de roca desprendida de la costa que se inclinaba peligrosamente sobre los escollos |
| but i sleep more comfortable in it without the blanket sheets mostly without the blanket | pero permítanme la de que me <UNK> <UNK> <UNK> la comisión | 0.18 | pero duermo más cómodo sin la manta de damasco y las sábanas bordadas en su mayoría sin la manta | 0.81 | Pero yo dormiría más a gusto en ella sin la colcha de damasco y las sábanas bordadas, principalmente sin la colcha |
| then why do you want the good things but for occasions such as this | entonces qué <UNK> de las comisión pero lo que puedo decir esta | 0.42 | entonces, ¿por qué quieres las cosas buenas, pero para ocasiones como esta | 1.00 | Pues ¿para qué se quieren las cosas buenas sino para las ocasiones como la presente |
| the quilt will be and how much it interferes with her in the bedroom | el ministro puede ser usted y con usted en la comisión | 0.20 | la colcha será y cuánto interfiere con ella en el dormitorio | 0.53 | Irá fuera la colcha y cuanto te estorbe con ella en la alcoba |
| here a doormat for your feet | en la un comisión le pido tus que <UNK> | 0.24 | aquí hay un felpudo para tus pies | 0.86 | Aquí tienes un felpudo para los pies |
| i think no evil will come to you at bedtime because these old chestnut soils are cold as they themselves | que no la <UNK> pero si no lo veo como <UNK> <UNK> | 0.15 | Creo que ningún mal vendrá a ti a la hora de dormir porque estos viejos suelos de castaño son fríos como ellos mismos | 0.84 | Creo que no te vendrá mal al acostarte, porque estos suelos de castaño viejo son fríos como ellos solos |
| as we are mortal and no free of trouble and the nights are so long now and so obscure and so cold and mostly you do not know | mientras <UNK> y <UNK> <UNK> y nos nos dice y las digo | 0.17 | como somos mortales y no estamos libres de problemas y las noches son tan largas y tan oscuras y tan frías y, sobre todo, no sabes | 0.87 | Como somos mortales y nadie está libre de un apuro, y las noches son tan largas ahora, y los carrejos tan obscuros y tan fríos y no los conoces tú mayormente |
| and in the opposite empty cupboard which you can put a clothes shop | y en si <UNK> la región señor barroso | 0.20 | y en el armario vacío opuesto que puedes poner una tienda de ropa | 0.56 | y en la de enfrente ese armario desocupado, en que puedes meter una tienda de ropa |
| that much you bring between him with leftovers you have to fit | esto es difícil que usted le y <UNK> | 0.15 | tanto que traes entre él y la cómoda y las perchas con los restos que tienes que caber | 0.88 | que por mucha que traigas, entre él y la cómoda y las perchas, con sobras te ha de caber |
| for your prayers some use it as a good christian you are to go to bed and leaving her behold at the head of our Lord on the cross and next with its water | por escrito bg de esta noche en chipre de la unión monetaria <UNK> o <UNK> <UNK> y | 0.24 | para sus oraciones algunos lo usan como un buen cristiano, deben irse a la cama y dejarla contemplar a la cabeza de nuestro Señor en la cruz y benditera, junto con su agua que corresponde a la bahía de ramuco, y bendecido si quieren rociar la habitación porque el el diablo no descansa un punto y se desliza por el ojo de un candado | 0.94 | Para tus rezos, porque alguno usarás, como buen cristiano que eres, al meterte en la cama y al salir de ella, ahí tienes, a la cabecera, a Dios Nuestro Señor en cruz, y la benditera al lado, con su agua correspondiente, y su ramuco de laurel bendito, por si quieres rociarla por el cuarto; porque el demonio no descansa un punto, y se cuela por el ojo de una cerradura |
| and still plenty of room for another much more | Aún y hay más de la situación actual | 0.15 | y todavía hay mucho espacio para otro mucho más | 1.00 | y todavía sobra campo para otro tanto más |
| what you trouble out with it if something you want and not have it ask for it which as you have at hand yours will be | mucho más bien de nuevo o de quién quiero decir la comisión y lo que puedo decir | 0.28 | lo que le moleste si algo quiere y no lo tiene, pídalo que, como lo tiene a mano, será | 0.86 | Lo que te estorbe, fuera con ello; si algo deseas y no lo tienes, pídelo, que, como lo haya a mano, tuyo será |
| and now i leave you in peace and your widths | y y le dejo deseo paz y <UNK> | 0.18 | Y ahora te dejo en paz y tus anchos | 0.70 | Y ahora te dejo en paz y a tus anchuras |
| when finished warns that we are in the kitchen | cuándo se <UNK> en el reino unido | 0.37 | cuando termine, advierte que estamos en la cocina | 0.75 | Cuando acabes, avisa, que en la cocina estamos |
| and he was shaking in one hand and requiring the other shelter that is slipping shoulder but a lot and breathing very anxious | ha sido una historia de historia y <UNK> y lo digo de todo corazón | 0.21 | Y sacudía la linterna con una mano y exigía que el otro refugio se deslizara por el hombro pero tosiera mucho y respirara muy ansioso. | 0.96 | Y se fue, zarandeando el farol en una mano y requiriendo con la otra el abrigo que se le deslizaba de los hombros; pero tosiendo mucho y muy anheloso de respiración |
| the railroad was not completely finished more than to the right place | la asamblea no ha respondido nada más que la cuestión | 0.19 | el ferrocarril no estaba completamente terminado más que en el lugar correcto | 0.78 | El camino de hierro no estaba completamente terminado mas que hasta el sitio indicado |
| today comes to turin the capital of piedmont which i had dreamed of beautiful is a beautiful city it is undeniable but extremely monotonous | hoy la historia de una unión europea rc <UNK> es un <UNK> y <UNK> | 0.26 | hoy viene a turin la capital de Piamonte que había soñado con hermosa es una hermosa ciudad es innegable pero extremadamente monótona | 0.92 | Hoy llega á Turin. #TURIN. # La capital del Piamonte, que yo habia soñado bellísima, es una ciudad muy linda, no puede negarse, pero extremadamente monótona é insustancial |
| all streets are cut from the same pattern annoying and equality so tired studied its magnificent which have been prudently distributed to the city appearance hold the view that limited circle monotonous tired to disgust | todos los refugiados nos <UNK> <UNK> <UNK> <UNK> <UNK> y <UNK> <UNK> <UNK> <UNK> | 0.13 | todas las calles están cortadas del mismo patrón molesto e igualdad tan cansado estudiado sus magníficos arcos que se han distribuido prudentemente a la apariencia de la ciudad embellecido mantienen la vista que círculo limitado monótono cansado a disgusto | 1.00 | Todas sus calles están cortadas por el mismo patron; fastidia y cansa una igualdad tan estudiada: sus arcadas magníficas, que prudentemente repartidas hubieran dado fisonomía á la ciudad, hermoseándola, encierran la vista en círculo tan limitado y monótono que cansan hasta el disgusto |
| although the town is small colossal abound the most beauty to me is the emmanuel to the extent the city | aunque el comunismo <UNK> me <UNK> y me gustaría compartir la palabra | 0.20 | aunque la ciudad es pequeña, abundan las plazas colosales, la mayor belleza para mí es la victorio emmanuel en la medida en que la ciudad | 1.00 | Aunque la ciudad es pequeña abundan las colosales plazas: la mas bella para mí es la de Victorio Emmanuel, al extremo de la ciudad |
| its extension is prodigious to find encircled by three walls of essential like all high and spacious | su intervención de <UNK> <UNK> <UNK> <UNK> y <UNK> | 0.20 | su extensión es prodigiosa de encontrar rodeada por tres paredes de arcos esenciales como todos altos y espaciosos | 0.82 | Su extension es prodigiosa; hállase circuida por tres fachadas de indispensables arcos, elevados y espaciosos como todos |
| opens in the side of the long street which leads straight to the square | <UNK> la situación de la situación y de la comisión | 0.37 | se abre en un semicírculo en el lado de la calle larga que conduce directamente a la plaza del palazzo vecchio | 0.91 | Ciérrase en forma de por el lado de la larga calle que conduce en línea recta á la plaza del Palazzo Vecchio |
| at the foot of the vast square is the historical tendency with its clear and abundant waves | la historia de la ue y la unión europea es totalmente eficiente y eficiente | 0.07 | al pie de la gran plaza se encuentra la tendencia histórica con sus olas claras y abundantes | 0.83 | A los piés de la inmensa plaza se tiende el histórico Pó, con sus cristalinas y abundantes ondas |
| from the center of that square is a table very picturesque the mighty river which saves an elegant stone bridge of five at the end of which and facing stands the church with its crown and a humble by colonnaded front right and left and closing the horizon artistic mountains pompously dressed vegetables | hasta la fecha de la historia de europa y francia <UNK> <UNK> <UNK> <UNK> y <UNK> y <UNK> <UNK> | 0.14 | desde el centro de esa plaza hay una mesa muy pintoresca, el poderoso río que salva un elegante puente de piedra de cinco arcos al final del cual se encuentra la iglesia de la gran madre di he con sus cúpulas y corona esbelta y un humilde peristilo de columnata delantera derecha e izquierda y cerrando el horizonte montañas artísticas verduras pomposamente vestidas | 0.95 | Desde el centro de la citada plaza se ofrece un cuadro muy pintoresco: el caudaloso rio, que salva un elegante puente de piedra de cinco arcos, al fin del cual, y dando frente, se levanta la iglesia de la Gran Madre di Dio, con su esbelta cupulita por corona, y un humilde peristilo con columnatas por frente: á derecha é izquierda y cerrando el horizonte, artísticas montañas, vestidas pomposamente de verdura |
| on the right and located on a scenic hill is presented a convent left the beautiful church of the elegant columns two towers | sobre el día de la unión soviética <UNK> <UNK> <UNK> <UNK> <UNK> | 0.23 | A la derecha y situado en una pintoresca colina se presenta un convento de frailes a la izquierda la hermosa iglesia de la superga con elegantes columnas dos torres y una elegante cúpula | 1.00 | A la derecha, y situado sobre una pintoresca colina, se presenta un convento de frailes; á la izquierda, la bellísima iglesia de la Superga, con elegantes columnas, dos torres y una airosa cúpula |
| before going further i want to point me the impression that was the view | antes de terminar permítanme que me gustaría destacar la misma opinión | 0.26 | Antes de ir más lejos quiero señalarme la impresión de que era la vista de los frailes | 1.00 | Antes de pasar adelante quiero apuntar la impresion que me produjo la vista de los frailes |
| in venice where he first was able to examine closely a monk | en el consejo <UNK> <UNK> <UNK> <UNK> | 0.14 | en Venecia donde fue capaz de examinar de cerca a un monje | 0.85 | En Venecia fué donde por primera vez pude examinar de cerca un monje |
| accustomed to hearing about them from my childhood in the makes sense now persuaded me that it is an institution happened fortunately for humanity and what happened never to return ever without having a clear idea of because in spain were abolished when i was seven i frankly confess that he passed to me the first to be offered to my view venice i left an almost spontaneously repulsive | hablar de una vez más allá de una reunión de ayer en septiembre de berlín dije que | 0.23 | acostumbrado a escuchar sobre ellos desde mi niñez en el sentido tiene ahora persuadido de que se trata de una institución afortunada para la humanidad y de lo que nunca volvió sin tener una idea clara de porque en España fueron abolidos cuando tenía siete años, francamente confieso que me pasó a mí el primero en ser ofrecido a mi punto de vista Venecia me dejó una respuesta casi espontáneamente repulsiva | 0.95 | Acostumbrado á oir hablar de ellos desde mi niñez, en el sentido que hoy se hace; persuadido yo de que es una institucion que pasó para fortuna de la humanidad, y que pasó para no volver jamas; sin tener una idea precisa de ellos, porque en España fueron abolidos cuando yo tenia siete años, confieso francamente que al pasar junto á mí el primero que se ofreció á mi vista en Venecia, me separé espontáneamente por un movimiento casi repulsivo |
| i watched him curiously short and when i lost sight of me congratulated again as he had done before that there i have in my country | he escuchado que usted le he escuchado en vano y quiero decir que en todos ellos | 0.16 | Lo vi curiosamente bajo y cuando me perdí de vista felicité de nuevo como lo había hecho antes de que haya en mi país | 0.93 | Le observé con curiosidad suma, y cuando le perdí de vista, me felicité nuevamente, como lo habia hecho otras veces, de que no les hubiera en mi pais |
| i am convinced beyond all doubt that the monks have passed forever and fortunately all human institutions have the time | alguien que todos los irlandeses <UNK> y <UNK> lo he hecho y lo haré | 0.16 | Estoy convencido, más allá de toda duda, de que los monjes han pasado para siempre y, afortunadamente, todas las instituciones humanas tienen el tiempo | 0.91 | Estoy hasta la evidencia convencido de que los monjes han pasado para siempre y por fortuna; las instituciones humanas todas tienen su época. |
| another good places turin is the wide spacious with majestic buildings | otro ejemplo es el de la <UNK> <UNK> <UNK> | 0.16 | Otro buen lugar turin es el carignano amplio espacioso con edificios majestuosos | 1.00 | Otra de las buenas plazas de Turin es la de Carignano: ancha y espaciosa, con majestuosos edificios |
| the old palace square behind which lies the king of sardinia in a building called the palace by mistake without doubt is large and clear as any in the city to it lead to the most elegant streets where trading has its luxury shops | el último homenaje homenaje a la historia de europa y la unión europea nos preocupa y <UNK> y <UNK> | 0.24 | la antigua plaza del palacio detrás de la cual se encuentra el rey de Cerdeña en un edificio llamado el palacio por error sin duda es grande y clara, ya que cualquiera en la ciudad conduce a las calles más elegantes donde el comercio tiene sus tiendas de lujo | 0.96 | La plaza del Palacio Viejo, detras del cual reside el rey de Cerdeña en un edificio que llaman palacio, por equivocacion sin duda, es grande y despejada como todas las de la ciudad: á ella desembocan las calles mas elegantes, donde el comercio tiene sus lujosas tiendas |
| the old palace which provides isolated in the middle of the square not interested one minute achieved in spite of his venerable age historical colour | el muro de silencio de chipre <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> | 0.11 | el antiguo palacio que proporciona aislado en el medio de la plaza no le interesa un minuto logrado a pesar de su edad venerable color histórico | 0.92 | El Palacio Viejo, que aislado se ofrece en medio de la plaza, no logró interesarme un minuto á pesar de su antigüedad venerable y su color histórico |
| his canvases uneven and abandoned their old towers its dark practicable balcony poor your wretched stature that does not even come with a lot to the size of the palaces all together in a word offer nothing to admire reason why we move forward and truth is a no small trouble to me i do not know where to take the reader as turin venice is different | <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> o <UNK> y <UNK> but arms lo digo de todas las versiones lingüísticas | 0.16 | sus lienzos irregulares y abandonaron sus viejas torres su oscuro balcón practicable pobre su miserable estatura que ni siquiera viene con mucho al tamaño de los palacios, todos juntos en una palabra no ofrecen nada que admirar por lo que avanzamos y la verdad es una no pequeña problema para mí, no sé dónde llevar al lector, ya que Turín Venecia es diferente | 1.00 | Sus lienzos desiguales y abandonados, sus viejos torreones, su balconaje oscuro y pobre, su mezquina estatura que no llega ni con mucho á la talla de los palacios, todo su conjunto en una palabra no ofrece nada que admirar, razon por la que seguimos adelante, y á la verdad que es un apuro no pequeño para mí, que no sé donde llevar al lector: ¡tan diferente es Turin de Venecia |
| one of the most beautiful capital of piedmont me of course and nothing is to men | una de las personas más o como <UNK> quiero hablar de la unión europea | 0.14 | uno de los ornamentos más bellos de la capital de Piamonte me, por supuesto, y nada es para los hombres | 0.95 | Uno de los mas bellos ornamentos de la capital del Piamonte para mí, es natural, y nada debe á los hombres |
| i speak of the magnificent alps always wonderfully impressive with its hair secular | he escuchado con mucho éxito con los <UNK> de la comisión | 0.29 | hablo de los magníficos Alpes siempre maravillosamente impresionante con su pelo blanco como la nieve secular | 0.93 | Hablo de los magníficos Alpes, siempre portentosamente admirables, con su secular cabellera de blanquísima nieve |
| the name of the nation amply indicates the situation of alps in turin piedmont at the foot of the mountains | la historia de la historia de los estados unidos lo que nos preocupa a los ciudadanos | 0.22 | el nombre de la nación indica ampliamente la situación de los Alpes en Turín Piamonte al pie de las montañas | 0.91 | El nombre de la nacion indica sobradamente la situacion de los Alpes en Turin: Piamonte, es decir, al pié de los montes |

**Appendix B**

**Sentence Translation from Machine Translator with Epoch 20 and Google Translator for Ferrero Dataset**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **What mattered was that he knew what so many times he had thought vaguely in the isolation of the tower without being able to fulfill his wishes** | qué ha ocurrido que lo que se ha hecho en el caso de que se le ha dicho que no se ha atrevido a la autoridad palestina | 0.37 | Lo que importaba era que supiera lo que tantas veces había pensado vagamente en el aislamiento de la torre sin poder cumplir con sus deseos | 0.88 | Lo importante era que conociese lo que tantas veces había pensado él vagamente en el aislamiento de la torre, sin poder dar forma precisa a sus deseos |
| **He continued his way slowly without reaching the family** | la comisión no puede responder a la familia familiar y familiar | 0.33 | siguió su camino lentamente sin llegar a la familia | 0.50 | Continuó lentamente su camino, para no alcanzar a la familia de Can |
| **he joined his mother and brother** | su presidente aplaude y su nombre de vigo | 0.33 | se unió a su madre y hermano | 0.66 | se había reunido con su madre y su hermano |
| **observed from a height when the group now walked towards the valley towards the farm** | <UNK> de la ciudad que se encuentra en el mismo año que se <UNK> en nuestro grupo de <UNK> | 0.16 | observó desde una altura cuando el grupo caminó ahora hacia el valle hacia la granja | 0.66 | Los vio desde una altura, cuando el grupo caminaba ya por el valle con dirección a la alquería |
| **twisted went to the pirate's tower, but when I got closer, he continued his way without stopping until the marcamino avoiding the approach** | de este modo la comisión puede utilizar su uso de la a energía | 0.28 | retorció fue a la torre del pirata, pero cuando me acerqué, siguió su camino sin detenerse hasta el marcamino evitando el enfoque | 0.42 | torció su marcha, evitando aproximarse a Can |
| **He went to the pirate's tower, but when I got close, he continued his way without stopping to the sea** | el sr brittan ha sido <UNK> pero no me parece que se <UNK> en el reino unido | 0.25 | fue a la torre del pirata, pero cuando me acerqué, siguió su camino sin detenerse hasta el mar | 1.00 | Fue hacia la torre del Pirata, pero al llegar cerca de ella continuó su camino, no deteniéndose hasta el mar |
| **this secular work was nibbling the coast snatching its stone shell sheet by leaf** | Bg la de por <UNK> <UNK> <UNK> <UNK> | 0.13 | este trabajo secular estaba mordisqueando la costa arrebatando su hoja de concha de piedra por hoja | 0.86 | Esta labor secular iba royendo la costa, arrebatándola su coraza de piedra, lámina por lámina |
| **fragments as walls** | en primer lugar la frase de la sra <UNK> <UNK> <UNK> <UNK> <UNK> | 0.33 | fragmentos lo como muros | 0.50 | de la ella fragmentos enormes como murallas |
| **first forming an imperceptible slit that got bigger with the course always** | me ha sorprendido a la vista gorda de la <UNK> de una especie de <UNK> <UNK> <UNK> | 0.28 | se primero formando una hendidur imperceptible que se hizo más grande con el curso siempre | 0.85 | primeramente formando una rendija imperceptible, que se agrandaba con el curso de los siglos |
| **he was to sit on the edge of a large rock he advanced a fragment of rock separated from** | fue un acto de una guerra de <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> | 0.10 | debía sentarse en el borde de una gran roca avanzó un fragmento de roca separada de | 0.74 | fue a sentarse en el borde de un gran peñasco avanzado, de un fragmento de roca desprendida de la costa que se inclinaba peligrosamente sobre los escollos |
| **but i sleep more comfortable in it without the blanket sheets mostly without the blanket** | pero no me parece más que esperar más y que la <UNK> de la <UNK> <UNK> | 0.38 | pero duermo más cómodo sin la manta de damasco y las sábanas bordadas en su mayoría sin la manta | 0.81 | Pero yo dormiría más a gusto en ella sin la colcha de damasco y las sábanas bordadas, principalmente sin la colcha |
| **then why do you want the good things but for occasions such as this** | por qué no lo hacen lo que se necesita es la respuesta de la comisión esta | 0.46 | entonces, ¿por qué quieres las cosas buenas, pero para ocasiones como esta | 1.00 | Pues ¿para qué se quieren las cosas buenas sino para las ocasiones como la presente |
| **the quilt will be and how much it interferes with her in the bedroom** | la comisión y el personal de la <UNK> se <UNK> de forma responsable | 0.30 | la colcha será y cuánto interfiere con ella en el dormitorio | 0.53 | Irá fuera la colcha y cuanto te estorbe con ella en la alcoba |
| **here a doormat for your feet** | en nombre un de la comisión <UNK> <UNK> <UNK> | 0.42 | aquí hay un felpudo para tus pies | 0.86 | Aquí tienes un felpudo para los pies |
| **i think no evil will come to you at bedtime because these old chestnut soils are cold as they themselves** | creo que no se puede utilizar en la <UNK> de la <UNK> de la que se nos <UNK> a nosotros mismos | 0.36 | Creo que ningún mal vendrá a ti a la hora de dormir porque estos viejos suelos de castaño son fríos como ellos mismos | 0.84 | Creo que no te vendrá mal al acostarte, porque estos suelos de castaño viejo son fríos como ellos solos |
| **as we are mortal and no free of trouble and the nights are so long now and so obscure and so cold and mostly you do not know** | como no estamos hablando de la guerra y la falta de dinero y de la palabra y no nos <UNK> y que no se le ha concedido | 0.26 | como somos mortales y no estamos libres de problemas y las noches son tan largas y tan oscuras y tan frías y, sobre todo, no sabes | 0.87 | Como somos mortales y nadie está libre de un apuro, y las noches son tan largas ahora, y los carrejos tan obscuros y tan fríos y no los conoces tú mayormente |
| **and in the opposite empty cupboard which you can put a clothes shop** | y en la que se puede decir que la <UNK> de la <UNK> <UNK> | 0.43 | y en el armario vacío opuesto que puedes poner una tienda de ropa | 0.56 | y en la de enfrente ese armario desocupado, en que puedes meter una tienda de ropa |
| **that much you bring between him with leftovers you have to fit** | de hecho le estoy más familiarizada que se le <UNK> a la comisión | 0.28 | tanto que traes entre él y la cómoda y las perchas con los restos que tienes que caber | 0.88 | que por mucha que traigas, entre él y la cómoda y las perchas, con sobras te ha de caber |
| **for your prayers some use it as a good christian you are to go to bed and leaving her behold at the head of our Lord on the cross and next with its water** | por su señoría que usted ha dicho que es un <UNK> de la unión y de la <UNK> de la <UNK> y de la leche desnatada a la vez que no se puede hacer nada en el caso de la palabra y en segundo lugar que no se trata de una <UNK> | 0.48 | para sus oraciones algunos lo usan como un buen cristiano, deben irse a la cama y dejarla contemplar a la cabeza de nuestro Señor en la cruz y benditera, junto con su agua que corresponde a la bahía de ramuco, y bendecido si quieren rociar la habitación porque el el diablo no descansa un punto y se desliza por el ojo de un candado | 0.94 | Para tus rezos, porque alguno usarás, como buen cristiano que eres, al meterte en la cama y al salir de ella, ahí tienes, a la cabecera, a Dios Nuestro Señor en cruz, y la benditera al lado, con su agua correspondiente, y su ramuco de laurel bendito, por si quieres rociarla por el cuarto; porque el demonio no descansa un punto, y se cuela por el ojo de una cerradura |
| **and still plenty of room for another much more** | y no tengo nada más hay que para mucho decir con la <UNK> | 0.50 | y todavía hay mucho espacio para otro mucho más | 1.00 | y todavía sobra campo para otro tanto más |
| **what you trouble out with it if something you want and not have it ask for it which as you have at hand yours will be** | lo que le les pide y que no se puede hacer y que no se utilice a la comisión y a la que no podemos utilizar | 0.34 | lo que le moleste si algo quiere y no lo tiene, pídalo que, como lo tiene a mano, será | 0.86 | Lo que te estorbe, fuera con ello; si algo deseas y no lo tienes, pídelo, que, como lo haya a mano, tuyo será |
| **and now i leave you in peace and your widths** | y ahora estoy de acuerdo con israel y herzegovina y herzegovina | 0.50 | Y ahora te dejo en paz y tus anchos | 0.70 | Y ahora te dejo en paz y a tus anchuras |
| **when finished warns that we are in the kitchen** | cuando se trata de la palabra en el <UNK> | 0.50 | cuando termine, advierte que estamos en la cocina | 0.75 | Cuando acabes, avisa, que en la cocina estamos |
| **and he was shaking in one hand and requiring the other shelter that is slipping shoulder but a lot and breathing very anxious** | y por lo tanto se ha hecho una <UNK> de la <UNK> y que se ha hecho pero me parece muy poco tiempo y <UNK> | 0.27 | Y sacudía la linterna con una mano y exigía que el otro refugio se deslizara por el hombro pero tosiera mucho y respirara muy ansioso. | 0.96 | Y se fue, zarandeando el farol en una mano y requiriendo con la otra el abrigo que se le deslizaba de los hombros; pero tosiendo mucho y muy anheloso de respiración |
| **the railroad was not completely finished more than to the right place** | no se ha llegado a mí la idea de que se en haya dicho | 0.21 | el ferrocarril no estaba completamente terminado más que en el lugar correcto | 0.78 | El camino de hierro no estaba completamente terminado mas que hasta el sitio indicado |
| **today comes to turin the capital of piedmont which i had dreamed of beautiful is a beautiful city it is undeniable but extremely monotonous** | hoy se ha hecho una reacción de la amenaza de la mantequilla de la mantequilla de la mantequilla que es una persona que es intolerable y <UNK> | 0.27 | hoy viene a turin la capital de Piamonte que había soñado con hermosa es una hermosa ciudad es innegable pero extremadamente monótona | 0.92 | Hoy llega á Turin. #TURIN.# La capital del Piamonte, que yo habia soñado bellísima, es una ciudad muy linda, no puede negarse, pero extremadamente monótona é insustancial |
| **all streets are cut from the same pattern annoying and equality so tired studied its magnificent which have been prudently distributed to the city appearance hold the view that limited circle monotonous tired to disgust** | todos los musulmanes se han visto afectados y que la unión europea se <UNK> y <UNK> en la que se <UNK> la pena de muerte de la <UNK> de los alimentos que se <UNK> de la <UNK> | 0.32 | todas las calles están cortadas del mismo patrón molesto e igualdad tan cansado estudiado sus magníficos arcos que se han distribuido prudentemente a la apariencia de la ciudad embellecido mantienen la vista que círculo limitado monótono cansado a disgusto | 1.00 | Todas sus calles están cortadas por el mismo patron; fastidia y cansa una igualdad tan estudiada: sus arcadas magníficas, que prudentemente repartidas hubieran dado fisonomía á la ciudad, hermoseándola, encierran la vista en círculo tan limitado y monótono que cansan hasta el disgusto |
| **although the town is small colossal abound the most beauty to me is the emmanuel to the extent the city** | aunque el <UNK> es un <UNK> de <UNK> <UNK><UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> | 0.125 | aunque la ciudad es pequeña, abundan las plazas colosales, la mayor belleza para mí es la victorio emmanuel en la medida en que la ciudad | 1.00 | Aunque la ciudad es pequeña abundan las colosales plazas: la mas bella para mí es la de Victorio Emmanuel, al extremo de la ciudad |
| **its extension is prodigious to find encircled by three walls of essential like all high and spacious** | su uso de la biomasa es que se <UNK> a todos los elementos de la comisión y 2013 | 0.32 | su extensión es prodigiosa de encontrar rodeada por tres paredes de arcos esenciales como todos altos y espaciosos | 0.82 | Su extension es prodigiosa; hállase circuida por tres fachadas de indispensables arcos, elevados y espaciosos como todos |
| **opens in the side of the long street which leads straight to the square** | <UNK> de la situación de la biomasa de la biomasa de todos los recursos | 0.28 | se abre en un semicírculo en el lado de la calle larga que conduce directamente a la plaza del palazzo vecchio | 0.91 | Ciérrase en forma de por el lado de la larga calle que conduce en línea recta á la plaza del Palazzo Vecchio |
| **at the foot of the vast square is the historical tendency with its clear and abundant waves** | en el caso de la amistad de la ue es el de la <UNK> y la <UNK> <UNK> | 0.33 | al pie de la gran plaza se encuentra la tendencia histórica con sus olas claras y abundantes | 0.83 | A los piés de la inmensa plaza se tiende el histórico Pó, con sus cristalinas y abundantes ondas |
| **from the center of that square is a table very picturesque the mighty river which saves an elegant stone bridge of five at the end of which and facing stands the church with its crown and a humble by colonnaded front right and left and closing the horizon artistic mountains pompously dressed vegetables** | desde el punto de vista de la planta <UNK> que se <UNK> a los <UNK> <UNK> y <UNK> <UNK> <UNK> y <UNK> a la <UNK> de la <UNK> y a la <UNK> de la <UNK> y a la <UNK> de los <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> | 0.11 | desde el centro de esa plaza hay una mesa muy pintoresca, el poderoso río que salva un elegante puente de piedra de cinco arcos al final del cual se encuentra la iglesia de la gran madre di he con sus cúpulas y corona esbelta y un humilde peristilo de columnata delantera derecha e izquierda y cerrando el horizonte montañas artísticas verduras pomposamente vestidas | 0.95 | Desde el centro de la citada plaza se ofrece un cuadro muy pintoresco: el caudaloso rio, que salva un elegante puente de piedra de cinco arcos, al fin del cual, y dando frente, se levanta la iglesia de la Gran Madre di Dio, con su esbelta cupulita por corona, y un humilde peristilo con columnatas por frente: á derecha é izquierda y cerrando el horizonte, artísticas montañas, vestidas pomposamente de verdura |
| **on the right and located on a scenic hill is presented a convent left the beautiful church of the elegant columns two towers** | en relación con la <UNK> y la <UNK> de la ue se <UNK> a una lista de <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> <UNK> | 0.19 | A la derecha y situado en una pintoresca colina se presenta un convento de frailes a la izquierda la hermosa iglesia de la superga con elegantes columnas dos torres y una elegante cúpula | 1.00 | A la derecha, y situado sobre una pintoresca colina, se presenta un convento de frailes; á la izquierda, la bellísima iglesia de la Superga, con elegantes columnas, dos torres y una airosa cúpula |
| **before going further i want to point me the impression that was the view** | antes de que me gustaría hacer una observación oral sobre la base de la palabra | 0.56 | Antes de ir más lejos quiero señalarme la impresión de que era la vista de los frailes | 1.00 | Antes de pasar adelante quiero apuntar la impresion que me produjo la vista de los frailes |
| **in venice where he first was able to examine closely a monk** | en el caso de la comisión ha de ser <UNK> en el año 2000 | 0.30 | en Venecia donde fue capaz de examinar de cerca a un monje | 0.85 | En Venecia fué donde por primera vez pude examinar de cerca un monje |
| **accustomed to hearing about them from my childhood in the makes sense now persuaded me that it is an institution happened fortunately for humanity and what happened never to return ever without having a clear idea of because in spain were abolished when i was seven i frankly confess that he passed to me the first to be offered to my view venice i left an almost spontaneously repulsive** | la reacción de la ue no ha sido capaz de que mi país se <UNK> en la que se <UNK> a mi país y <UNK> pero creo que es una <UNK> que nos ha dicho que la comisión no ha hecho un <UNK> de <UNK> <UNK> | 0.27 | acostumbrado a escuchar sobre ellos desde mi niñez en el sentido tiene ahora persuadido de que se trata de una institución afortunada para la humanidad y de lo que nunca volvió sin tener una idea clara de porque en España fueron abolidos cuando tenía siete años, francamente confieso que me pasó a mí el primero en ser ofrecido a mi punto de vista Venecia me dejó una respuesta casi espontáneamente repulsiva | 0.95 | Acostumbrado á oir hablar de ellos desde mi niñez, en el sentido que hoy se hace; persuadido yo de que es una institucion que pasó para fortuna de la humanidad, y que pasó para no volver jamas; sin tener una idea precisa de ellos, porque en España fueron abolidos cuando yo tenia siete años, confieso francamente que al pasar junto á mí el primero que se ofreció á mi vista en Venecia, me separé espontáneamente por un movimiento casi repulsivo |
| **i watched him curiously short and when i lost sight of me congratulated again as he had done before that there i have in my country** | yo he escuchado con la palabra y que se ha hecho en mi intervención por lo que se ha hecho en el pasado y yo diría que no se repita | 0.25 | Lo vi curiosamente bajo y cuando me perdí de vista felicité de nuevo como lo había hecho antes de que haya en mi país | 0.93 | Le observé con curiosidad suma, y cuando le perdí de vista, me felicité nuevamente, como lo habia hecho otras veces, de que no les hubiera en mi pais |
| **i am convinced beyond all doubt that the monks have passed forever and fortunately all human institutions have the time** | estoy seguro de que todos los que se han hecho en el mundo y en la que todos los medios de comunicación se <UNK> | 0.31 | Estoy convencido, más allá de toda duda, de que los monjes han pasado para siempre y, afortunadamente, todas las instituciones humanas tienen el tiempo | 0.91 | Estoy hasta la evidencia convencido de que los monjes han pasado para siempre y por fortuna; las instituciones humanas todas tienen su época. |
| **another good places turin is the wide spacious with majestic buildings** | otra es muy bien la política de pesca de bajura de los edificios | 0.35 | Otro buen lugar turin es el carignano amplio espacioso con edificios majestuosos | 1.00 | Otra de las buenas plazas de Turin es la de Carignano: ancha y espaciosa, con majestuosos edificios |
| **the old palace square behind which lies the king of sardinia in a building called the palace by mistake without doubt is large and clear as any in the city to it lead to the most elegant streets where trading has its luxury shops** | la campaña de la campaña de la enfermedad de alzheimer y de la <UNK> de la sociedad es muy grave y que la <UNK> se <UNK> en la que se nos ha dado la espalda de la ue y no se puede utilizar en el mundo | 0.21 | la antigua plaza del palacio detrás de la cual se encuentra el rey de Cerdeña en un edificio llamado el palacio por error sin duda es grande y clara, ya que cualquiera en la ciudad conduce a las calles más elegantes donde el comercio tiene sus tiendas de lujo | 0.96 | La plaza del Palacio Viejo, detras del cual reside el rey de Cerdeña en un edificio que llaman palacio, por equivocacion sin duda, es grande y despejada como todas las de la ciudad: á ella desembocan las calles mas elegantes, donde el comercio tiene sus lujosas tiendas |
| **the old palace which provides isolated in the middle of the square not interested one minute achieved in spite of his venerable age historical colour** | el <UNK> de la <UNK> de la <UNK> de la vida de los niños no se puede en cualquier caso de <UNK> de la vida familiar | 0.33 | el antiguo palacio que proporciona aislado en el medio de la plaza no le interesa un minuto logrado a pesar de su edad venerable color histórico | 0.92 | El Palacio Viejo, que aislado se ofrece en medio de la plaza, no logró interesarme un minuto á pesar de su antigüedad venerable y su color histórico |
| **his canvases uneven and abandoned their old towers its dark practicable balcony poor your wretched stature that does not even come with a lot to the size of the palaces all together in a word offer nothing to admire reason why we move forward and truth is a no small trouble to me i do not know where to take the reader as turin venice is different** | su esposa de su marido y de la historia de la unión europea no es un <UNK> y <UNK> y que no se trata de una forma de <UNK> pero que nos parece que es muy importante que la unión europea se convierta en una especie de <UNK> | 0.43 | sus lienzos irregulares y abandonaron sus viejas torres su oscuro balcón practicable pobre su miserable estatura que ni siquiera viene con mucho al tamaño de los palacios, todos juntos en una palabra no ofrecen nada que admirar por lo que avanzamos y la verdad es una no pequeña problema para mí, no sé dónde llevar al lector, ya que Turín Venecia es diferente | 1.00 | Sus lienzos desiguales y abandonados, sus viejos torreones, su balconaje oscuro y pobre, su mezquina estatura que no llega ni con mucho á la talla de los palacios, todo su conjunto en una palabra no ofrece nada que admirar, razon por la que seguimos adelante, y á la verdad que es un apuro no pequeño para mí, que no sé donde llevar al lector: ¡tan diferente es Turin de Venecia |
| **one of the most beautiful capital of piedmont me of course and nothing is to men** | una de las formas más bajas de la familia no es la de la <UNK> y la solidaridad | 0.33 | uno de los ornamentos más bellos de la capital de Piamonte me, por supuesto, y nada es para los hombres | 0.95 | Uno de los mas bellos ornamentos de la capital del Piamonte para mí, es natural, y nada debe á los hombres |
| **i speak of the magnificent alps always wonderfully impressive with its hair secular** | me he abstenido en la idea de que todos los presos de la unión europea <UNK> | 0.27 | hablo de los magníficos Alpes siempre maravillosamente impresionante con su pelo blanco como la nieve secular | 0.93 | Hablo de los magníficos Alpes, siempre portentosamente admirables, con su secular cabellera de blanquísima nieve |
| **the name of the nation amply indicates the situation of alps in turin piedmont at the foot of the mountains** | la idea de que se <UNK> la palabra de la presidencia británica de la unión europea en nombre de los verdes | 0.36 | el nombre de la nación indica ampliamente la situación de los Alpes en Turín Piamonte al pie de las montañas | 0.81 | El nombre de la nacion indica sobradamente la situacion de los Alpes en Turin: Piamonte, es decir, al pié de los montes |

**Appendix C**

**Sentence Translation from Machine Translator with Epoch 10 and Google Translator for UN Dataset**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Indeed it is quite in keeping with the positions this House has always adopted | la decisión de acuerdo con otros oradores no ha presentado ninguna decisión | 0.31 | De hecho, está en consonancia con las posiciones que esta Cámara siempre ha adoptado | 0.68 | Esta sin duda dentro de la linea de las posiciones que nuestro Parlamento siempre ha adoptado |
| Why are there no fire instructions | por qué no hay ninguna objeción | 0.44 | ¿Por qué no hay instrucciones de fuego? | 0.77 | ?Por que no hay instrucciones en caso de incendio |
| I would also like to say that the Commission is prepared and ready to hold this debate whenever it is convenient and that we were ready to do so this week as we had agreed originally on the basis that it would be presented the day before in a speech to parliamentary groups | señor comisario usted también en la votación de la comisión de agricultura he votado en contra y pido que la comisión y demás decisiones | 0.21 | También me gustaría decir que la Comisión está preparada y lista para celebrar este debate cuando sea conveniente y que estábamos listos para hacerlo esta semana, ya que habíamos acordado originalmente sobre la base de que se presentaría el día anterior en un discurso. a grupos parlamentarios | 0.84 | Y quiero decir senora Presidenta que por parte de la Comision estamos preparados y dispuestos a tener ese debate cuando convenga que estabamos preparados para desarrollarlo esta semana como en principio se habia acordado partiendo de la base de que se presentaba la vispera en un discurso a los grupos parlamentarios |
| I propose that we vote on the request of the Group of the Party of European Socialists that the Commission statement on its strategic objectives should be reinstated | la comisión de presupuestos de la comisión de presupuestos de este parlamento me gustaría aclarar la subsidiariedad | 0.26 | Propongo que votemos a petición del Grupo del Partido de los Socialistas Europeos de que se restablezca la declaración de la Comisión sobre sus objetivos estratégicos | 0.92 | Propongo que votemos la peticion del Grupo del Partido de los Socialistas Europeos para volver a incluir la declaracion de la Comision sobre sus objetivos estrategicos |
| All of this is in accordance with the principle of subsidiarity and is therefore to be greatly welcomed | nada de la comisión de peticiones y subsidiariedad también debe ser rechazado por la ponente | 0.50 | Todo esto está de acuerdo con el principio de subsidiariedad y, por lo tanto, debe ser muy bien recibido | 0.42 | Hay que celebrar extraordinariamente todo esto en el sentido de la subsidiariedad |
| The report looks at the issue of harmonising the examination requirements for safety advisors working in the areas of transportation of dangerous goods by road rail and inland waterway | el informe <UNK> el aumento de la fiebre <UNK> y <UNK> <UNK> <UNK> y <UNK> | 0.13 | El informe analiza la cuestión de la armonización de los requisitos de examen para los asesores de seguridad que trabajan en las áreas de transporte de mercancías peligrosas por ferrocarril y vía navegable | 0.86 | El informe trata de la cuestion de armonizar los requisitos de examen para los consejeros de seguridad que desempenan su tarea en el ambito del transporte de mercancias peligrosas por carretera vias ferreas y vias fluviales |
| This is a pity in a sense | es una cuestión muy importante | 0.33 | Esto es una pena en cierto sentido | 1.00 | Es una pena en cierto sentido |
| But I believe that we should do all we can to keep the transport of dangerous goods to a minimum in all countries whether they are transit countries or not | pero tengo que todo todo de todos los países sino que <UNK> o no <UNK> | 0.27 | Pero creo que debemos hacer todo lo posible para mantener el transporte de mercancías peligrosas al mínimo en todos los países, ya sean países de tránsito o no | 0.93 | Pero creo que todos debemos hacer todo lo posible para restringir el transporte de mercancias peligrosas y por cierto en todos los paises sean paises de transito o no |
| Very briefly I would like to thank the various Members for their interventions and to tell you that safety is one of the Commission s priorities in the field of transport | más felicidades felicidades por la comisión de presupuestos de polonia en la comisión de calidad o en otro modo | 0.23 | Muy brevemente, me gustaría dar las gracias a los diversos miembros por sus intervenciones y decirles que la seguridad es una de las prioridades de la Comisión en el ámbito del transporte | 0.90 | Brevisimamente quiero agradecer tambien la intervencion de los distintos diputados y decirles a sus Senorias que la seguridad es una prioridad de la Comision en el ambito de los transportes |
| Transport of dangerous goods by road | transporte de mercancías peligrosas | 0.50 | Transporte de mercancías peligrosas por carretera | 1.00 | Transporte de mercancias peligrosas por carretera |
| we cannot be satisfied with just this however as after this target has been met there will still be too many areas where pollution counts exceed levels that are dangerous to people and nature | nadie nadie nadie ayer por esta situación de acuerdo con otros desastres y no podemos y debemos preocuparnos | 0.20 | no podemos estar satisfechos solo con esto, sin embargo, una vez alcanzado este objetivo, todavía habrá demasiadas áreas donde el recuento de contaminantes excederá los niveles que son peligrosos para las personas y la naturaleza | 0.92 | Sin embargo, no podemos conformarnos con esto porque, una vez alcanzado este objetivo, seguirán existiendo demasiadas zonas que sobrepasen en tal medida los valores límite de contaminantes, que continuarán siendo peligrosas para las personas y para el medio ambiente |
| owing to this i have recommended in my report a long term goal in addition to the interim one and a final target | he leído que he leído la última conclusión y una importante conclusión de la comisión y el final | 0.29 | Debido a esto, he recomendado en mi informe un objetivo a largo plazo además del provisional y un objetivo final | 0.83 | Por ello, además del objetivo intermedio, he propuesto en mi informe un objetivo a largo plazo y un objetivo definitivo para el año 2020. |
| in that way we would finally be able to reach a point where critical levels and loads are not exceeded and everyone will be protected effectively from all known health effects of air pollution | de este modo <UNK> <UNK> <UNK> <UNK> <UNK> de la unión y de la subsidiariedad de la subsidiariedad y el principio | 0.24 | de esa forma, finalmente podríamos llegar a un punto en el que los niveles críticos y las cargas no se excedan y todos estarán protegidos de forma efectiva contra todos los efectos conocidos de la contaminación del aire | 0.87 | Para entonces podríamos, por fin, llegar a una situación en la que los niveles e índices críticos no fuesen rebasados y todas las personas se verían protegidas eficazmente frente a los conocidos efectos nocivos para la salud de los contaminantes atmosféricos. |
| in this connection i would like to draw your attention to my amendment in which owing to a translation error there is a wrong word in your copy of the document | en este sentido me he referido a una situación en este sentido en este punto quiero llamar la pregunta y en esto | 0.30 | a este respecto, quisiera llamar su atención sobre mi enmienda, en la que, debido a un error de traducción, hay una palabra incorrecta en su copia del documento. | 0.84 | En este contexto, quiero hacer una observación sobre este aspecto de mi enmienda, pues, debido a un fallo lingüístico, en sus documentos hay una palabra incorrecta. |
| in place of reached it should say exceeded | en de realidad debería <UNK> la comisión | 0.25 | en lugar de alcanzarlo debería decir excedido | 0.75 | En lugar de la palabra alcanzado hay que poner la palabra rebasado. |
| i hope that this will be sorted out before tomorrow vote | espero que se <UNK> mañana a la votación | 0.38 | Espero que esto se solucione antes de mañana | 0.92 | Espero que esto sea tomado en cuenta antes de la votación de mañana. |
| how then do we achieve a situation in which air pollution does not exceed tolerance levels for humans and for nature | qué pasa en este conflicto en el mercado laboral no podemos y también la subsidiariedad | 0.26 | ¿cómo podemos lograr una situación en la que la contaminación del aire no exceda los niveles de tolerancia para los humanos y para la naturaleza? | 0.92 | Entonces, ¿cómo se podrá llegar a una situación en que los contaminantes atmosféricos no sobrepasen los límites de tolerancia de la naturaleza y de las personas? |
| the commission proposal sets limit values for each member state with regard to certain kinds of interrelated emissions which cause acidification increased ozone in the inner atmosphere and soil eutrophication | la propuesta de la comisión de la unión monetaria <UNK> <UNK> <UNK> <UNK> <UNK> y <UNK> en inglés | 0.16 | la propuesta de la comisión establece valores límite para cada estado miembro con respecto a ciertos tipos de emisiones interrelacionadas que causan acidificación, aumento del ozono en la atmósfera interna y eutroficación del suelo | 0.89 | En la propuesta de la Comisión se establecen para cada Estado miembro límites máximos de emisión de determinados contaminantes, relacionados entre sí, que provocan acidificación, aumentan el índice de ozono en las capas bajas de la atmósfera y dan lugar a una fertilización excesiva del suelo. |
| dealing with these problems and the emissions that cause them together has proved to be the most cost effective way of tackling the causes of air pollution and bringing it under control | ninguna de estas contradicciones de estos problemas y los problemas de la eeb nos hemos hecho y también de la trazabilidad | 0.38 | El tratamiento de estos problemas y las emisiones que los causan juntos ha demostrado ser la forma más rentable de abordar las causas de la contaminación del aire y ponerla bajo control | 0.93 | El examen conjunto de estos problemas y de los contaminantes que los originan se ha revelado como un método económicamente eficaz para descubrir los causantes de la contaminación y poder controlarlos. |
| the biggest disagreements in the implementation of the directive relate to the limit values the emission ceiling set for each country | la mayoría de las negociaciones de adhesión y los criterios de copenhague nos permiten aprender | 0.20 | los mayores desacuerdos en la aplicación de la directiva se relacionan con los valores límite establecidos para cada país | 0.54 | Las mayores discrepancias que se plantean en la consecución de esta directiva están relacionadas precisamente con el nivel de estos límites máximos, es decir, con los valores objetivo de contaminación atribuidos a cada Estado miembro. |
| the europe has also examined the objectives for reductions in the same sources of emissions as in the proposal for a directive now under discussion and as a result of these talks the so called gothenburg protocol was signed | el parlamento ha hecho una ampliación de la unión monetaria y de otras semanas <UNK> y también en el consejo | 0.15 | Europa también ha examinado los objetivos de reducción de las mismas fuentes de emisiones que en la propuesta de directiva que ahora se está debatiendo y como resultado de estas conversaciones se firmó el llamado protocolo de Gothenburg | 0.96 | En el marco de la Comisión Económica para Europa, que depende de la ONU, también se han examinado los objetivos de reducción de las mismas emisiones contaminantes que las de esta propuesta de directiva que ahora nos ocupa, y como resultado de esos debates se ha decidido firmar el llamado Protocolo de Goteburgo. |
| there is a clear difference between this proposal and the commission | hay una propuesta de la comisión y también al ponente | 0.53 | hay una clara diferencia entre esta propuesta y la comisión | 1.00 | Hay una clara diferencia entre esta propuesta y la propuesta de la Comisión. |
| if we were to accept the ceilings in the gothenburg protocol we would never reach the targets that the union has set itself regarding air pollution | si la comisión de presupuestos de este tipo de situaciones también se aplica la sostenibilidad de la directiva | 0.28 | si tuviéramos que aceptar los límites del protocolo de Gotemburgo, nunca alcanzaríamos los objetivos que el sindicato se ha fijado en relación con la contaminación del aire | 0.75 | Si nos conformamos con el nivel de valores del Protocolo de Goteburgo, nunca llegaremos a alcanzar los objetivos que la Unión Europea se ha fijado a sí misma en cuanto a la tasa de contaminantes atmosféricos. |
| comparing the differences between the gothenburg protocol and the commission proposal with regard to the effects on health the annual reduction levels proposed by the commission would save the lives of as many as four thousand people a year in europe. | <UNK> la comisión ni entre las negociaciones de negociación o en 1996 en grecia y deberíamos decir de la salud | 0.15 | Al comparar las diferencias entre el protocolo de Gotemburgo y la propuesta de la comisión con respecto a los efectos sobre la salud, los niveles anuales de reducción propuestos por la comisión salvarían las vidas de hasta cuatro mil personas al año en Europa | 0.93 | Si comparamos las diferencias existentes entre el Protocolo de Goteburgo y la propuesta de la Comisión desde el punto de vista de las consecuencias para la salud, el nivel de reducción propuesto por la Comisión puede llegar a salvar incluso cuarenta mil vidas humanas al año en Europa. |
| although it is beyond dispute that the ambitious targets set for the environment have been found to be an essential ingredient in people health and safety there have been pleas for lower ceilings on the basis that cleaner air costs too much for industry and communities | aunque se ha tratado de acuerdo en este marco de empleo life de empleo life y hortalizas y también los temores y la comisión | 0.23 | Aunque es indiscutible que los ambiciosos objetivos para el medioambiente se han encontrado como un ingrediente esencial en la salud y seguridad de las personas, se han pedido límites más bajos sobre la base de que un aire más limpio cuesta demasiado para la industria y las comunidades | 0.84 | A pesar de que se ha demostrado irrefutablemente que los objetivos ambiciosos en materia medioambiental contribuyen a la salud y a la seguridad de los ciudadanos, se ha defendido el establecimiento de valores objetivo más bajos argumentando que un aire más limpio resulta demasiado caro para la industria y para la sociedad. |
| it has to be said however that there is very considerable exaggeration with regard to the costs involved | hay otra observación que hay que analizar las cosas que se plantean a la comisión | 0.60 | Sin embargo, debe decirse que existe una exageración muy considerable con respecto a los costos involucrados | 0.86 | Sin embargo, hay que decir que se han exagerado demasiado los costes en este sentido. |
| if on the other hand we calculate the benefits that better air quality will bring to people and the natural environment as well as to buildings we see that the economic benefits are four times greater han the costs | pero <UNK> con el mundo de los animales de los animales pero también la cuestión de los costes y de la comisión | 0.16 | si, por otro lado, calculamos los beneficios que una mejor calidad del aire traerá a las personas y al medio ambiente natural, así como a los edificios, vemos que los beneficios económicos son cuatro veces mayores que los costos | 0.76 | Mientras que si ciframos las ventajas que se desprenden de la existencia de una mejor calidad del aire tanto para los ciudadanos, como para la naturaleza y para las construcciones, sacamos el resultado de que los beneficios económicos son cuatro veces mayores que los costes. |
| we still have to work on the proposal for the directive so that for example the situation with regard to the applicant countries and those questions too that relate to the neighbouring regions can be better taken into account | hemos llegado a una hora de trabajo en 1996 y grecia en grecia deberíamos considerar otras cosas y de ello | 0.20 | todavía tenemos que trabajar en la propuesta de Directiva para que, por ejemplo, la situación con respecto a los países candidatos y las preguntas que también se relacionan con las regiones vecinas se puedan tener mejor en cuenta | 0.88 | Tenemos que ser capaces de moldear aún esta propuesta de directiva de tal modo que se tome mejor en cuenta, por ejemplo, la situación de los países candidatos a la adhesión, así como las cuestiones relacionadas con los territorios vecinos de la Unión Europea. |
| mr president the industry committee after much discussion and serious consideration is overwhelmingly opposed to the commission proposed ceilings and this is across all groups and nationalities | señor presidente la comisión de trabajo y de mujer y mujeres todos los hombres y los demás | 0.27 | señor presidente del comité de la industria después de mucha discusión y seria consideración se opone abrumadoramente a los techos propuestos por la comisión y esto es en todos los grupos y nacionalidades | 0.94 | Señor Presidente, la Comisión de Industria, después de mucho debate y de un examen a fondo, se opone de manera arrolladora a los límites propuestos por la Comisión, y ello en todos los grupos y nacionalidades. |
| the ceilings were agreed by all our member states by the applicant countries and globally in a protocol signed at gothenburg as recently as november 1999 | el parlamento decidió de turquía y junio de junio de junio también la comisión y el 26 de mayo | 0.21 | los techos fueron acordados por todos nuestros estados miembros por los países candidatos y a nivel mundial en un protocolo firmado en Gothenburg en fecha tan reciente como noviembre de 1999 | 0.93 | Los límites de la CEPE han sido aprobados por todos nuestros Estados miembros, por los países candidatos y, a escala mundial, en un protocolo firmado en Goteburgo tan recientemente como noviembre de 1999. |
| why did the commission come up with these even stricter and substantially more expensive ceilings only one month later | por último la comisión decidió y <UNK> <UNK> <UNK> <UNK> | 0.22 | ¿Por qué la comisión propuso estos techos aún más estrictos y sustancialmente más caros solo un mes después? | 0.88 | ¿Por qué propuso la Comisión estos límites aún más estrictos y sustancialmente más caros sólo un mes después? |
| i understand that the groups the whole parliament are split on this issue between supporters of the committee on the environment public health and consumer policy and supporters of the industry committee line | mis señorías en la unión europea en pie de la comisión de empleo y turismo comprendo que la comisión y el informe | 0.33 | Entiendo que los grupos de todo el parlamento están divididos en este tema entre los partidarios del comité sobre la salud pública del medio ambiente y la política del consumidor y los partidarios de la línea del comité de la industria | 0,94 | Entiendo que los grupos, todo el Parlamento, estén divididos acerca de este tema, entre los partidarios de la Comisión de Medio Ambiente, Salud Pública y Política del Consumidor y los partidarios de la línea de la Comisión de Industria. |
| members of parliament should be aware of the serious implications for our industries and agriculture for our competitiveness for jobs and for people and our constituents if the commission proposal wins through | alguien de la comisión de desarrollo rural en esta unión deberíamos concentrar otra pregunta ayuda contra la subsidiariedad y la opinión | 0.20 | los miembros del parlamento deben ser conscientes de las graves implicaciones para nuestras industrias y la agricultura de nuestra competitividad para el empleo y para las personas y nuestros electores si la propuesta de la comisión gana a través de | 0.88 | Los diputados al Parlamento debieran ser conscientes de las graves consecuencias para nuestras industrias y nuestro sector agrario, para nuestra capacidad competitiva, para los puestos de trabajo y -cómo no- para la gente y para los habitantes de nuestras circunscripciones electorales si al fin triunfa la propuesta de la Comisión. |
| the industry committee urges this house to vote in favour of the amendments which are tabled in my name and in the name of mr bowe | la comisión europea rechazó una propuesta pública y de todas las enmiendas señor verheugen y a favor | 0.30 | el comité de la industria insta a esta cámara a votar a favor de las enmiendas que se presentan a mi nombre y en nombre de mr bowe | 0.92 | La Comisión de Industria insta a la Asamblea a votar a favor de las enmiendas que se presentan en mi nombre y en el del Sr. Bowe. |
| these are the amendments passed by the industry committee | son las enmiendas de la comisión de industria | 0.66 | estas son las enmiendas aprobadas por el comité de la industria | 0.88 | Son las enmiendas aprobadas por la Comisión de Industria. |
| the un ceilings are quite ambitious enough for us to reach our internationally agreed targets in reasonable time | los estados unidos nos <UNK> de nuevo de los fondos para el informe de todos modos | 0.38 | los límites máximos de la ONU son lo suficientemente ambiciosos como para que podamos alcanzar nuestros objetivos internacionalmente acordados en un tiempo razonable | 0.76 | Los límites de las Naciones Unidas son bastante ambiciosos como para que alcancemos nuestras metas aceptadas internacionalmente en un plazo razonable. |
| the industry committee wishes to see these limits fully and properly implemented across the european union and globally | la comisión de presupuestos europeos ha apoyado y de hecho y también de la confianza | 0.40 | el comité de industria desea ver estos límites completa y adecuadamente implementados en toda la unión europea y globalmente | 0.81 | La Comisión de Industria desea que estos límites se apliquen completa y adecuadamente en toda la Unión Europea y a escala mundial. |
| the commission can then concentrate on ensuring this occurs on perfecting the models that are used and perhaps come up with new proposals in two or three years time on the basis of much more convincing arguments regarding the relevant costs and benefits | la comisión presentará a la vez más allá de acuerdo con la comisión de dos criterios y criterios y criterios y queremos que todos | 0.31 | la comisión puede entonces concentrarse en asegurar que esto ocurra al perfeccionar los modelos que se usan y tal vez proponer nuevas propuestas dentro de dos o tres años sobre la base de argumentos mucho más convincentes con respecto a los costos y beneficios relevantes | 0.90 | La Comisión puede entonces centrarse en velar por esto ocurra, en perfeccionar los modelos que se aplican y sacar quizá nuevas propuestas dentro de dos o tres años basándose en unos argumentos mucho más convincentes sobre los costes y beneficios conexos. |
| support the purvis amendments and the united nations gothenburg table of national emission ceilings | las enmiendas de los dos parlamentos y los ministros de finanzas nacionales también deben aplicarse | 0.38 | apoyar las enmiendas purvis y la tabla de techos nacionales de emisión de las naciones unidas de Gotemburgo | 0.94 | Apoyen las enmiendas Purvis y el cuadro de límites nacionales de emisión de las Naciones Unidas de Goteburgo. |
| mr president i wish to begin by thanking mrs myller for her report and her ambitious treatment of the directive proposed by the commission | señor presidente quería terminar felicitando a la primera lectura de la pregunta de la comisión y al informe | 0.38 | señor presidente, quiero comenzar dando las gracias a la señora Myller por su informe y su ambicioso tratamiento de la directiva propuesta por la comisión | 0.93 | Señor Presidente, quiero empezar dando las gracias a la Sra. Myller por su informe y por el dinámico tratamiento que ha dado a la directiva presentada por la Comisión. |
| even if the committee on industry external trade research and energy appears rather as if it had been caught napping and has only just woken up the fact may at all events be conveyed that the problem of acidification is scarcely a new issue instead it has been with us for as long as industrialisation for 100 years | incluso la comisión de presupuestos de 1996 y <UNK> <UNK> <UNK> <UNK> <UNK> y <UNK> a todos los años | 0.32 | incluso si el comité sobre investigación y energía del comercio exterior de la industria parece más bien como si hubiera sido sorprendido y acabara de despertar, en cualquier caso puede transmitirse que el problema de la acidificación no es un problema nuevo, sino que ha estado con nosotros. durante el tiempo que la industrialización durante 100 años | 0.92 | Aunque parece que la Comisión de Industria, Comercio Exterior, Investigación y Energía ha despertado súbitamente y estaba un poco distraída, podemos informar que los problemas de la acidificación no son nuevos; han existido tanto tiempo como el industrialismo, cien años. |
| research into the question has been going on for 30 years | la propuesta de la comisión ha sido un 30 de años | 0.62 | la investigación de la pregunta ha estado sucediendo durante 30 años | 0.75 | Esta materia se ha investigado durante 30 años. |
| none of us ought to be in any doubt about what the sources of acidification are nor about what we should be doing regarding this issue | nada de nadie de acuerdo con el problema de la cuestión deberíamos reflexionar también de esto | 0.28 | ninguno de nosotros debería tener ninguna duda sobre cuáles son las fuentes de acidificación ni sobre lo que deberíamos hacer con respecto a este tema | 0.94 | Ninguno de nosotros debería tener dudas sobre las fuentes de la acidificación y las medidas para combatirla. |
| some accuse the commission proposal and also the report as being excessive and far too costly to implement and say that the calculations that have been made are utopian | algunas propuestas de la comisión de presupuestos y de acuerdo con <UNK> lo que también quería decir y lo que se lo he | 0.56 | algunos acusan a la propuesta de la comisión y también al informe de ser excesivo y demasiado costoso de implementar y decir que los cálculos que se han hecho son utópicos | 0.92 | Hay quienes sostienen que la propuesta de la Comisión y el informe son exagerados y demasiado costosos, además, dicen que los cálculos son utópicos. |
| this is however wrong because we know that the member states themselves have contributed to the report which serves as a basis for the whole directive | pero no obstante hemos votado en contra de este tipo de responsabilidad que el consejo tiene que ver | 0.32 | sin embargo, esto es incorrecto porque sabemos que los propios Estados miembros han contribuido al informe que sirve de base para toda la directiva | 0.92 | Sin embargo, es erróneo decir esto, ya que sabemos que los propios países miembros han proporcionado los datos en los que se fundamenta la directiva |

**Appendix D**

**Sentence Translation from Machine Translator with Epoch 20 and Google Translator for UN Dataset**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Indeed it is quite in keeping with the positions this House has always adopted | de hecho es totalmente inaceptable que se haya votado en contra de la votación nominal | 0.43 | De hecho, está en consonancia con las posiciones que esta Cámara siempre ha adoptado | 0.68 | Esta sin duda dentro de la linea de las posiciones que nuestro Parlamento siempre ha adoptado |
| Why are there no fire instructions | por qué no se hay trata de la enmienda <UNK> | 0.55 | ¿Por qué no hay instrucciones de fuego? | 0.77 | ?Por que no hay instrucciones en caso de incendio |
| I would also like to say that the Commission is prepared and ready to hold this debate whenever it is convenient and that we were ready to do so this week as we had agreed originally on the basis that it would be presented the day before in a speech to parliamentary groups | también me gustaría decir que la decisión que vamos a votar y que se apruebe y yo creo que se ha presentado y que lo que se ha dicho en el pleno en la votación de hoy hemos pedido a la comisión de peticiones que se han pedido a todos los grupos | 0.78 | También me gustaría decir que la Comisión está preparada y lista para celebrar este debate cuando sea conveniente y que estábamos listos para hacerlo esta semana, ya que habíamos acordado originalmente sobre la base de que se presentaría el día anterior en un discurso. a grupos parlamentarios | 0.84 | Y quiero decir senora Presidenta que por parte de la Comision estamos preparados y dispuestos a tener ese debate cuando convenga que estabamos preparados para desarrollarlo esta semana como en principio se habia acordado partiendo de la base de que se presentaba la vispera en un discurso a los grupos parlamentarios |
| I propose that we vote on the request of the Group of the Party of European Socialists that the Commission statement on its strategic objectives should be reinstated | propongo que votemos a favor de la votación de la comisión de medio ambiente de la comisión de medio ambiente que se asegure de que se aprueben con el consejo | 0.88 | Propongo que votemos a petición del Grupo del Partido de los Socialistas Europeos de que se restablezca la declaración de la Comisión sobre sus objetivos estratégicos | 0.92 | Propongo que votemos la peticion del Grupo del Partido de los Socialistas Europeos para volver a incluir la declaracion de la Comision sobre sus objetivos estrategicos |
| All of this is in accordance with the principle of subsidiarity and is therefore to be greatly welcomed | todo esto es aceptable por lo que ser refiere a la tanto aprobación de la gestión y bien la trazabilidad | 0.72 | Todo esto está de acuerdo con el principio de subsidiariedad y, por lo tanto, debe ser muy bien recibido | 0.42 | Hay que celebrar extraordinariamente todo esto en el sentido de la subsidiariedad |
| The report looks at the issue of harmonising the examination requirements for safety advisors working in the areas of transportation of dangerous goods by road rail and inland waterway | el informe aborda la cuestión de la seguridad de los productos de los productos de seguridad en los productos de los productos de transporte de mercancías y servicios de terceros países | 0.80 | El informe analiza la cuestión de la armonización de los requisitos de examen para los asesores de seguridad que trabajan en las áreas de transporte de mercancías peligrosas por ferrocarril y vía navegable | 0.86 | El informe trata de la cuestion de armonizar los requisitos de examen para los consejeros de seguridad que desempenan su tarea en el ambito del transporte de mercancias peligrosas por carretera vias ferreas y vias fluviales |
| This is a pity in a sense | esto es una que en parece muy sentido | 0.83 | Esto es una pena en cierto sentido | 1.00 | Es una pena en cierto sentido |
| But I believe that we should do all we can to keep the transport of dangerous goods to a minimum in all countries whether they are transit countries or not | pero creo que deberíamos hacer todo lo posible para que se utilice la importación de productos de transporte de mercancías peligrosas o no que no se importen o no mercancías | 0.82 | Pero creo que debemos hacer todo lo posible para mantener el transporte de mercancías peligrosas al mínimo en todos los países, ya sean países de tránsito o no | 0.93 | Pero creo que todos debemos hacer todo lo posible para restringir el transporte de mercancias peligrosas y por cierto en todos los paises sean paises de transito o no |
| Very briefly I would like to thank the various Members for their interventions and to tell you that safety is one of the Commission s priorities in the field of transport | a mi juicio quiero dar las gracias a los distintos diputados y a la comisión que se trata de una cuestión de que se pueda utilizar en el ámbito de la investigación transporte | 0.63 | Muy brevemente, me gustaría dar las gracias a los diversos miembros por sus intervenciones y decirles que la seguridad es una de las prioridades de la Comisión en el ámbito del transporte | 0.90 | Brevisimamente quiero agradecer tambien la intervencion de los distintos diputados y decirles a sus Senorias que la seguridad es una prioridad de la Comision en el ambito de los transportes |
| Transport of dangerous goods by road | transporte de mercancías peligrosas por carbono | 1.00 | Transporte de mercancías peligrosas por carretera | 1.00 | Transporte de mercancias peligrosas por carretera |
| we cannot be satisfied with just this however as after this target has been met there will still be too many areas where pollution counts exceed levels that are dangerous to people and nature | no obstante no se puede decir que en este momento se <UNK> que en el caso de que se trata de una actividad de los recursos hídricos y otros factores son muy importantes y en la actualidad | 0.48 | no podemos estar satisfechos solo con esto, sin embargo, una vez alcanzado este objetivo, todavía habrá demasiadas áreas donde el recuento de contaminantes excederá los niveles que son peligrosos para las personas y la naturaleza | 0.92 | Sin embargo, no podemos conformarnos con esto porque, una vez alcanzado este objetivo, seguirán existiendo demasiadas zonas que sobrepasen en tal medida los valores límite de contaminantes, que continuarán siendo peligrosas para las personas y para el medio ambiente |
| owing to this i have recommended in my report a long term goal in addition to the interim one and a final target | a mi pesar a la conclusión de que se ha alcanzado un informe objetivo final en del marco de la revisión intermedia y definitiva | 0.57 | Debido a esto, he recomendado en mi informe un objetivo a largo plazo además del provisional y un objetivo final | 0.83 | Por ello, además del objetivo intermedio, he propuesto en mi informe un objetivo a largo plazo y un objetivo definitivo para el año 2020. |
| in that way we would finally be able to reach a point where critical levels and loads are not exceeded and everyone will be protected effectively from all known health effects of air pollution | de este modo podremos hacer un poco a la hora de que no se produzcan criterios y métodos de vida y que también se trata de la protección de los animales | 0.76 | de esa forma, finalmente podríamos llegar a un punto en el que los niveles críticos y las cargas no se excedan y todos estarán protegidos de forma efectiva contra todos los efectos conocidos de la contaminación del aire | 0.87 | Para entonces podríamos, por fin, llegar a una situación en la que los niveles e índices críticos no fuesen rebasados y todas las personas se verían protegidas eficazmente frente a los conocidos efectos nocivos para la salud de los contaminantes atmosféricos. |
| in this connection i would like to draw your attention to my amendment in which owing to a translation error there is a wrong word in your copy of the document | en este sentido quisiera llamar mi atención sobre la cuestión de la pena de muerte que se ha hecho una pregunta que se ha planteado en el libro verde | 0.75 | a este respecto, quisiera llamar su atención sobre mi enmienda, en la que, debido a un error de traducción, hay una palabra incorrecta en su copia del documento. | 0.84 | En este contexto, quiero hacer una observación sobre este aspecto de mi enmienda, pues, debido a un fallo lingüístico, en sus documentos hay una palabra incorrecta. |
| in place of reached it should say exceeded | de en hecho se podría decir que la navidad de excedido <UNK> | 0.41 | en lugar de alcanzarlo debería decir excedido | 0.75 | En lugar de la palabra alcanzado hay que poner la palabra rebasado. |
| i hope that this will be sorted out before tomorrow vote | espero que se aplace esta votación mañana a la votación nominal | 0.79 | Espero que esto se solucione antes de mañana | 0.92 | Espero que esto sea tomado en cuenta antes de la votación de mañana. |
| how then do we achieve a situation in which air pollution does not exceed tolerance levels for humans and for nature | cómo se puede hacer frente a la situación en que no se trata de una cuestión de contaminación y el agua potable | 0.80 | ¿cómo podemos lograr una situación en la que la contaminación del aire no exceda los niveles de tolerancia para los humanos y para la naturaleza? | 0.92 | Entonces, ¿cómo se podrá llegar a una situación en que los contaminantes atmosféricos no sobrepasen los límites de tolerancia de la naturaleza y de las personas? |
| the commission proposal sets limit values for each member state with regard to certain kinds of interrelated emissions which cause acidification increased ozone in the inner atmosphere and soil eutrophication | la propuesta de la comisión establece un nuevo acuerdo de valores que se <UNK> a los valores límite de los alimentos que se han producido en la lista de agua potable y <UNK> | 0.46 | la propuesta de la comisión establece valores límite para cada estado miembro con respecto a ciertos tipos de emisiones interrelacionadas que causan acidificación, aumento del ozono en la atmósfera interna y eutroficación del suelo | 0.89 | En la propuesta de la Comisión se establecen para cada Estado miembro límites máximos de emisión de determinados contaminantes, relacionados entre sí, que provocan acidificación, aumentan el índice de ozono en las capas bajas de la atmósfera y dan lugar a una fertilización excesiva del suelo. |
| dealing with these problems and the emissions that cause them together has proved to be the most cost effective way of tackling the causes of air pollution and bringing it under control | con estos problemas y los problemas que se han hecho son los que se han hecho la crisis de la contaminación atmosférica la contaminación atmosférica y la utilización de la carne de bovino | 0.51 | El tratamiento de estos problemas y las emisiones que los causan juntos ha demostrado ser la forma más rentable de abordar las causas de la contaminación del aire y ponerla bajo control | 0.93 | El examen conjunto de estos problemas y de los contaminantes que los originan se ha revelado como un método económicamente eficaz para descubrir los causantes de la contaminación y poder controlarlos. |
| the biggest disagreements in the implementation of the directive relate to the limit values the emission ceiling set for each country | las posiciones de la ue son los principios de la aplicación de los criterios de emisión de los biocarburantes de los biocarburantes | 0.37 | los mayores desacuerdos en la aplicación de la directiva se relacionan con los valores límite establecidos para cada país | 0.54 | Las mayores discrepancias que se plantean en la consecución de esta directiva están relacionadas precisamente con el nivel de estos límites máximos, es decir, con los valores objetivo de contaminación atribuidos a cada Estado miembro. |
| the europe has also examined the objectives for reductions in the same sources of emissions as in the proposal for a directive now under discussion and as a result of these talks the so called gothenburg protocol was signed | la ue ha hecho hincapié en la necesidad de establecer un paquete de medidas de la ocde en el marco de la propuesta de resolución y por lo tanto lo que se ha dicho en el anexo de julio de 2004 | 0.56 | Europa también ha examinado los objetivos de reducción de las mismas fuentes de emisiones que en la propuesta de directiva que ahora se está debatiendo y como resultado de estas conversaciones se firmó el llamado protocolo de Gothenburg | 0.96 | En el marco de la Comisión Económica para Europa, que depende de la ONU, también se han examinado los objetivos de reducción de las mismas emisiones contaminantes que las de esta propuesta de directiva que ahora nos ocupa, y como resultado de esos debates se ha decidido firmar el llamado Protocolo de Goteburgo. |
| there is a clear difference between this proposal and the commission | hay una clara diferencia de la propuesta de la comisión y el consejo | 0.60 | hay una clara diferencia entre esta propuesta y la comisión | 1.00 | Hay una clara diferencia entre esta propuesta y la propuesta de la Comisión. |
| if we were to accept the ceilings in the gothenburg protocol we would never reach the targets that the union has set itself regarding air pollution | si no se nos ha parecido en la cumbre de 1999 que se nos dice que se <UNK> el principio de cautela que nos ocupa | 0.36 | si tuviéramos que aceptar los límites del protocolo de Gotemburgo, nunca alcanzaríamos los objetivos que el sindicato se ha fijado en relación con la contaminación del aire | 0.75 | Si nos conformamos con el nivel de valores del Protocolo de Goteburgo, nunca llegaremos a alcanzar los objetivos que la Unión Europea se ha fijado a sí misma en cuanto a la tasa de contaminantes atmosféricos. |
| comparing the differences between the gothenburg protocol and the commission proposal with regard to the effects on health the annual reduction levels proposed by the commission would save the lives of as many as four thousand people a year in europe. | el hecho de que las negociaciones entre la comisión y el consejo sobre el empleo de la política de cohesión se <UNK> en la reducción de las cifras de la ue que se <UNK> en el año rural y en mi circunscripción | 0.47 | Al comparar las diferencias entre el protocolo de Gotemburgo y la propuesta de la comisión con respecto a los efectos sobre la salud, los niveles anuales de reducción propuestos por la comisión salvarían las vidas de hasta cuatro mil personas al año en Europa | 0.93 | Si comparamos las diferencias existentes entre el Protocolo de Goteburgo y la propuesta de la Comisión desde el punto de vista de las consecuencias para la salud, el nivel de reducción propuesto por la Comisión puede llegar a salvar incluso cuarenta mil vidas humanas al año en Europa. |
| although it is beyond dispute that the ambitious targets set for the environment have been found to be an essential ingredient in people health and safety there have been pleas for lower ceilings on the basis that cleaner air costs too much for industry and communities | aunque se han establecido importantes objetivos de protección de los alimentos de salud y medio ambiente y de la política agrícola se ha puesto que la comisión se centra en la conservación de los productos lácteos y otros que se utilizan en los otros países | 0.38 | Aunque es indiscutible que los ambiciosos objetivos para el medioambiente se han encontrado como un ingrediente esencial en la salud y seguridad de las personas, se han pedido límites más bajos sobre la base de que un aire más limpio cuesta demasiado para la industria y las comunidades | 0.84 | A pesar de que se ha demostrado irrefutablemente que los objetivos ambiciosos en materia medioambiental contribuyen a la salud y a la seguridad de los ciudadanos, se ha defendido el establecimiento de valores objetivo más bajos argumentando que un aire más limpio resulta demasiado caro para la industria y para la sociedad. |
| it has to be said however that there is very considerable exaggeration with regard to the costs involved | es sin cierto que hay que debe decir que hay que hacer frente a los costes que se han hecho | 0.76 | Sin embargo, debe decirse que existe una exageración muy considerable con respecto a los costos involucrados | 0.86 | Sin embargo, hay que decir que se han exagerado demasiado los costes en este sentido. |
| if on the other hand we calculate the benefits that better air quality will bring to people and the natural environment as well as to buildings we see that the economic benefits are four times greater than the costs | si a la vez que se les permite que se produzcan costes y costes de la industria alimentaria el consumo de la biodiversidad y los costes de los alimentos son los más importantes que nos afectan a los costes de los costes | 0.47 | si, por otro lado, calculamos los beneficios que una mejor calidad del aire traerá a las personas y al medio ambiente natural, así como a los edificios, vemos que los beneficios económicos son cuatro veces mayores que los costos | 0.76 | Mientras que si ciframos las ventajas que se desprenden de la existencia de una mejor calidad del aire tanto para los ciudadanos, como para la naturaleza y para las construcciones, sacamos el resultado de que los beneficios económicos son cuatro veces mayores que los costes. |
| we still have to work on the proposal for the directive so that for example the situation with regard to the applicant countries and those questions too that relate to the neighbouring regions can be better taken into account | seguimos de la directiva sobre la situación de la situación de la situación de la región y también en los estados miembros que se encuentran en las regiones ultraperiféricas y las regiones ultraperiféricas que se han planteado en la política | 0.36 | todavía tenemos que trabajar en la propuesta de Directiva para que, por ejemplo, la situación con respecto a los países candidatos y las preguntas que también se relacionan con las regiones vecinas se puedan tener mejor en cuenta | 0.88 | Tenemos que ser capaces de moldear aún esta propuesta de directiva de tal modo que se tome mejor en cuenta, por ejemplo, la situación de los países candidatos a la adhesión, así como las cuestiones relacionadas con los territorios vecinos de la Unión Europea. |
| mr president the industry committee after much discussion and serious consideration is overwhelmingly opposed to the commission proposed ceilings and this is across all groups and nationalities | señor presidente la comisión de asuntos económicos y turismo ha hecho un informe muy claro y que no se ha presentado en todos los aspectos de la omc y los programas | 0.83 | señor presidente del comité de la industria después de mucha discusión y seria consideración se opone abrumadoramente a los techos propuestos por la comisión y esto es en todos los grupos y nacionalidades | 0.94 | Señor Presidente, la Comisión de Industria, después de mucho debate y de un examen a fondo, se opone de manera arrolladora a los límites propuestos por la Comisión, y ello en todos los grupos y nacionalidades. |
| the ceilings were agreed by all our member states by the applicant countries and globally in a protocol signed at gothenburg as recently as november 1999 | los países candidatos han decidido cumplir todos los estados miembros de la ue y en el caso de las exportaciones de la comunidad en julio de 2004 y 14 de julio | 0.48 | los techos fueron acordados por todos nuestros estados miembros por los países candidatos y a nivel mundial en un protocolo firmado en Gothenburg en fecha tan reciente como noviembre de 1999 | 0.93 | Los límites de la CEPE han sido aprobados por todos nuestros Estados miembros, por los países candidatos y, a escala mundial, en un protocolo firmado en Goteburgo tan recientemente como noviembre de 1999. |
| why did the commission come up with these even stricter and substantially more expensive ceilings only one month later | por qué he visto que la comisión se estrictos y se caros con más o menos de la misma y de la omc | 0.84 | ¿Por qué la comisión propuso estos techos aún más estrictos y sustancialmente más caros solo un mes después? | 0.88 | ¿Por qué propuso la Comisión estos límites aún más estrictos y sustancialmente más caros sólo un mes después? |
| i understand that the groups the whole parliament are split on this issue between supporters of the committee on the environment public health and consumer policy and supporters of the industry committee line | comprendo que los grupos de interés de la comisión de asuntos de la comisión de medio ambiente salud pública y asuntos del grupo del partido del mercado interior y del empleo | 0.43 | Entiendo que los grupos de todo el parlamento están divididos en este tema entre los partidarios del comité sobre la salud pública del medio ambiente y la política del consumidor y los partidarios de la línea del comité de la industria | 0.94 | Entiendo que los grupos, todo el Parlamento, estén divididos acerca de este tema, entre los partidarios de la Comisión de Medio Ambiente, Salud Pública y Política del Consumidor y los partidarios de la línea de la Comisión de Industria. |
| members of parliament should be aware of the serious implications for our industries and agriculture for our competitiveness for jobs and for people and our constituents if the commission proposal wins through | los diputados de los diputados tienen que hacer frente a la gestión de nuestros mercados laborales y a la seguridad de nuestros ciudadanos y a la comisión de empleo y a la comisión de medio ambiente | 0.54 | los miembros del parlamento deben ser conscientes de las graves implicaciones para nuestras industrias y la agricultura de nuestra competitividad para el empleo y para las personas y nuestros electores si la propuesta de la comisión gana a través de | 0.88 | Los diputados al Parlamento debieran ser conscientes de las graves consecuencias para nuestras industrias y nuestro sector agrario, para nuestra capacidad competitiva, para los puestos de trabajo y -cómo no- para la gente y para los habitantes de nuestras circunscripciones electorales si al fin triunfa la propuesta de la Comisión. |
| the industry committee urges this house to vote in favour of the amendments which are tabled in my name and in the name of mr bowe | la comisión de industria peticiones se opone a la votación de la ponente que se ha presentado en mi grupo y a la sra mcnally | 0.65 | el comité de la industria insta a esta cámara a votar a favor de las enmiendas que se presentan a mi nombre y en nombre de mr bowe | 0.92 | La Comisión de Industria insta a la Asamblea a votar a favor de las enmiendas que se presentan en mi nombre y en el del Sr. Bowe. |
| these are the amendments passed by the industry committee | estas son las enmiendas presentadas por la comisión de medio ambiente industria | 0.95 | estas son las enmiendas aprobadas por el comité de la industria | 0.88 | Son las enmiendas aprobadas por la Comisión de Industria. |
| the un ceilings are quite ambitious enough for us to reach our internationally agreed targets in reasonable time | los compromisos internacionales están dispuestos a que se haga un calendario para todos los instrumentos que nos fijamos en el calendario | 0.45 | los límites máximos de la son lo suficientemente ambiciosos como para que podamos alcanzar nuestros objetivos internacionalmente acordados en un tiempo razonable | 0.76 | Los límites de las Naciones Unidas son bastante ambiciosos como para que alcancemos nuestras metas aceptadas internacionalmente en un plazo razonable. |
| the industry committee wishes to see these limits fully and properly implemented across the european union and globally | la comisión de industria de la ue se esfuerza a todos los instrumentos de pesca y de la union comunidad y de la comunidad | 0.53 | el comité de industria desea ver estos límites completa y adecuadamente implementados en toda la unión europea y globalmente | 0.81 | La Comisión de Industria desea que estos límites se apliquen completa y adecuadamente en toda la Unión Europea y a escala mundial. |
| the commission can then concentrate on ensuring this occurs on perfecting the models that are used and perhaps come up with new proposals in two or three years time on the basis of much more convincing arguments regarding the relevant costs and benefits | la comisión puede presentar a la comisión que se <UNK> a estas medidas y en segundo lugar si se <UNK> más y más tiempo y dos propuestas concretas sobre los costes de los recursos de los recursos y los costes de los recursos | 0.58 | la comisión puede entonces concentrarse en asegurar que esto ocurra al perfeccionar los modelos que se usan y tal vez proponer nuevas propuestas dentro de dos o tres años sobre la base de argumentos mucho más convincentes con respecto a los costos y beneficios relevantes | 0.90 | La Comisión puede entonces centrarse en velar por esto ocurra, en perfeccionar los modelos que se aplican y sacar quizá nuevas propuestas dentro de dos o tres años basándose en unos argumentos mucho más convincentes sobre los costes y beneficios conexos. |
| support the purvis amendments and the united nations gothenburg table of national emission ceilings | apoyo las enmiendas relativas a la table de resolución de la comisión y a unidays de Goteburgo los artículos | 0.72 | apoyar las enmiendas purvis y la tabla de techos nacionales de emisión de las naciones unidas de Gotemburgo | 0.94 | Apoyen las enmiendas Purvis y el cuadro de límites nacionales de emisión de las Naciones Unidas de Goteburgo. |
| mr president i wish to begin by thanking mrs myller for her report and her ambitious treatment of the directive proposed by the commission | señor presidente quería terminar felicitando a la primera lectura de la pregunta de la comisión y al informe | 0.35 | señor presidente, quiero comenzar dando las gracias a la señora Myller por su informe y su ambicioso tratamiento de la directiva propuesta por la comisión | 0.93 | Señor Presidente, quiero empezar dando las gracias a la Sra. Myller por su informe y por el dinámico tratamiento que ha dado a la directiva presentada por la Comisión. |
| even if the committee on industry external trade research and energy appears rather as if it had been caught napping and has only just woken up the fact may at all events be conveyed that the problem of acidification is scarcely a new issue instead it has been with us for as long as industrialisation for 100 years | aunque la comisión de medio ambiente y energía y el comercio de armas y la <UNK> de la unión europea se ha convertido parece bien en una cuestión muy importante que nos encontramos probolema nosotros en el segundo momento de que se nos <UNK> a los dos minutos de los que se nos ha concedido | 0.56 | incluso si el comité sobre investigación y energía del comercio exterior de la industria parece más bien como si hubiera sido sorprendido y acabara de despertar, en cualquier caso puede transmitirse que el problema de la acidificación no es un problema nuevo, sino que ha estado con nosotros. durante el tiempo que la industrialización durante 100 años | 0.92 | Aunque parece que la Comisión de Industria, Comercio Exterior, Investigación y Energía ha despertado súbitamente y estaba un poco distraída, podemos informar que los problemas de la acidificación no son nuevos; han existido tanto tiempo como el industrialismo, cien años. |
| research into the question has been going on for 30 years | se ha mencionado la pregunta durante atención estado sobre la duración de la durante gripe aviar | 0.62 | la investigación de la pregunta ha estado sucediendo durante 30 años | 0.75 | Esta materia se ha investigado durante 30 años. |
| none of us ought to be in any doubt about what the sources of acidification are nor about what we should be doing regarding this issue | ninguno de nosotros no podemos estar de acuerdo con el ni sobre deberiamos contenido de los daños de la contaminación y en este tema que no podemos hablar de la cuestión | 0.83 | ninguno de nosotros debería tener ninguna duda sobre cuáles son las fuentes de acidificación ni sobre lo que deberíamos hacer con respecto a este tema | 0.94 | Ninguno de nosotros debería tener dudas sobre las fuentes de la acidificación y las medidas para combatirla. |
| some accuse the commission proposal and also the report as being excessive and far too costly to implement and say that the calculations that have been made are utopian | algunas propuestas de la comisión y el informe que se ha propuesto y que se <UNK> y que se <UNK> la propuesta de la comisión y de la leche desnatada | 0.70 | algunos acusan a la propuesta de la comisión y también al informe de ser excesivo y demasiado costoso de implementar y decir que los cálculos que se han hecho son utópicos | 0.92 | Hay quienes sostienen que la propuesta de la Comisión y el informe son exagerados y demasiado costosos, además, dicen que los cálculos son utópicos. |
| this is however wrong because we know that the member states themselves have contributed to the report which serves as a basis for the whole directive | sin embargo esto es algo que nos preocupa que los estados miembros hayan votado a favor de la aprobación de la gestión de la directiva | 0.88 | sin embargo, esto es incorrecto porque sabemos que los propios Estados miembros han contribuido al informe que sirve de base para toda la directiva | 0.92 | Sin embargo, es erróneo decir esto, ya que sabemos que los propios países miembros han proporcionado los datos en los que se fundamenta la directiva |