



# Language rehabilitation of people with BROCA aphasia using deep neural machine translation

Kamel Smaïli, David Langlois, Peter Pribil

## ► To cite this version:

Kamel Smaïli, David Langlois, Peter Pribil. Language rehabilitation of people with BROCA aphasia using deep neural machine translation. Fifth International Conference on Computational Linguistics in Bulgaria, Sep 2022, Sofia, Bulgaria. hal-03701615

**HAL Id: hal-03701615**

**<https://hal.science/hal-03701615>**

Submitted on 22 Jun 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Language rehabilitation of people with BROCA aphasia using deep neural machine translation

Kamel Smaili

David Langlois

Peter Pribil

Loria - SMarT Université Lorraine

{kamel.smaili, david.langlois, peter.pribil}@loria.fr

## Abstract

More than 13 million people suffer a stroke each year. Aphasia is known as a language disorder usually caused by a stroke that damages a specific area of the brain that controls the expression and understanding of language. Aphasia is characterized by a disturbance of the linguistic code affecting encoding and/or decoding of the language. Our project aims to propose a method that helps a person suffering from aphasia to communicate better with those around him. For this, we will propose a machine translation capable of correcting aphasic errors and helping the patient to communicate more easily. To build such a system, we need a parallel corpus; to our knowledge, this corpus does not exist, especially for French. Therefore, the main challenge and the objective of this task is to build a parallel corpus composed of sentences with aphasic errors and their corresponding correction. We will show how we create a pseudo-aphasia corpus from real data, and then we will show the feasibility of our project to translate from aphasia data to natural language. The preliminary results show that the deep learning methods we used achieve correct translations corresponding to a BLEU of 38.6.

**Keywords:** Aphasia, Augmentation corpus, Machine translation, Deep learning.

## 1 Introduction

Stroke represents the 2<sup>nd</sup> cause of death in the population, and the 1<sup>st</sup> cause of physical handicap in France. According to World stroke organization, 13.7 million people worldwide will suffer their first stroke on 2022 and 5.5 million will die from it. The incidence of stroke increases significantly with age and, in the West, as people are living longer and longer, stroke is almost becoming a pandemic problem.

Stroke can affect the mechanisms of speech, movement, sensation, and so on. The physical, cognitive and psychic after-effects of a stroke remain,

unfortunately, frequent (30 to 50% of cases). Many survivors will experience some form of lifelong disability or impairment that they will attempt to cure. They are particularly confronted with the reduction of their movements and isolation, among other things, due to their inability to communicate normally. With rehabilitation and specialist support, however, most stroke survivors can return to a near-normal life.

According to the National Aphasia Association<sup>1</sup>, a third of strokes result in aphasia, a major after-effect that greatly affects quality of life (Summers et al., 2009). Aphasia, a term suggested by Armand Trousseau in 1863, is characterized by a disturbance of the linguistic code, affecting the expression and/or the comprehension, and which can concern the oral and/or the written language. It is a localized or diffuse brain damage, generally in the frontal, parietal and/or temporal area of the left hemisphere, essentially of vascular, traumatic or of tumor origin (Marshall et al., 1998).

There are several different types of Aphasia, all of them coming with their own unique side-effects (Clough and Gordon, 2020). Their classification is not a trivial task, however, there is one thing they all share: making communication a difficult task. Findings in current theory (Cho-Reyes and Thompson, 2012) suggest frequent misuses of verbs and nouns, either from a character-mismatch or lexical swap perspective, and heavy syntactic alterations (Garraffa and Fyndanis, 2020). The discourse abilities might also be limited (Armstrong, 2000).

Our ultimate goal is to help People with Aphasia (PwA) to find their words easily by offering them a speech-to-speech system that corrects mispronounced sentences. To achieve this, we first need a parallel corpus where the source is composed by the altered spoken sentences and the target by what should have been spoken. In our knowledge, this

<sup>1</sup><https://www.aphasia.org/aphasia-resources/aphasia-factsheet/>

kind of corpus does not exist. What we propose in this article is to **create such dataset by starting with sentences pronounced by PwA in speech therapy sessions and their correction, and then augment the corpus with sentence pairs automatically created based on the features of the initial data.** We will also perform some preliminary translation experiments to show the overall feasibility of the approach that will lead to an aphasic speech correction system.

Our focus is on Broca patients. We believe that given the nature of popular rehabilitation methods, such as linguistic specific treatment (LST) (Thompson et al., 2003) and mapping therapy (Rochon et al., 2005), both of which are based on repetition of words, similar structures, or giving clues on remembering certain words or phrases, our instant feedback system based on speech translation would be of great help.

## 2 Related works

In natural language processing (NLP) and also in humanities, the availability of corpora is essential for understanding behavior phenomena and proposing tools or softwares based on NLP techniques or machine learning methods that facilitate the comprehension of such phenomena. In this particular topic, the aphasia data are rare and those that exist are not available.

**One of the most attractive project developing Aphasia corpora is probably AphasiaBank** (Forbes et al., 2012). The objective of AphasiaBank project was to **provide researchers and clinicians with a large shared multimedia database of uniform discourse from individuals with and without aphasia.** The database includes language samples in English, Spanish, German, Italian, Hungarian, Mandarin and Chinese. The aphasia section of this database contains approximately 180 videos of people with aphasia.

The project RELEASE (Williams et al., 2016) refers to the aphasia dataset of individual patient data for the rehabilitation and recovery of people with Aphasia after stroke. This project seems to be used by clinicians with the objective to study the rehabilitation. No information is given about the transcription of their utterances.

In the Moss Aphasia Psycholinguistics Project (MAPP) (Mirman et al., 2010), the authors provide a searchable database with data from more than 240 patients. The database is made up of the

Philadelphia Naming Test (PNT) results. The PNT is a single-word picture naming test developed to collect a large corpus of naming answers from patients.

Concerning the works that have addressed the problem of aphasia using automatic language processing approaches, we can cite the research below.

Since the grammatical deficiencies depend on the Primary Progressive Aphasia (PPA), in (Themistocleous et al., 2021) the authors propose to classify PPA variants by using part of speech (POS) production and to identify morphological markers that classify them by using machine learning. PPA is a very unique kind of aphasia. It is a form of dementia, and there are no cures available. Eventually, the person with this dementia completely loses their ability to comprehend and produce language due to gradual degradation (Thompson et al., 1997).

The study (Day et al., 2021) combines natural language processing and machine learning methods to predict the severity of PwA, both by score and severity level. The authors used a dataset of 238 participants extracted from AphasiaBank. They took the data from its transcript and composed the dataset by removing stop words and other items not necessary for this task. Stop word lists differ greatly, but they usually contain non-thematic words, like function words (determiners), prepositions (on, it, under), and so on. This is a very questionable decision, given the importance of the already few words people with aphasia are uttering. Stop words could be important indicators.

## 3 Building an Aphasic-French parallel corpus

In this section, we will describe **how to build an Aphasic-French corpus (APHAFRECH)** which will be used to show the feasibility of developing a communication rehabilitation support system for an aphasic person. To do this, we **started by collecting real aphasia data in French that we transcribed**, then we **developed methods to build a parallel corpus that can be used to develop a machine translation system.** We used several sources to build up a corpus for the analysis of aphasic errors. The **first source** is made up of **videos extracted from the Web recorded in therapy sessions** between speech therapists and PwA. In each video **a speech therapist asked several questions** to the PwA such as: *What is your name? How do you feel today? Describe*

what you see in this picture. We transcribed the PwA utterance and we corrected it. We retrieved seven dialogues that last from 3 to 20 minutes, the statistics concerning these videos are given in Table 1.

$d$	65'8''
$ \bar{d} $	8'8''
$ \bar{w} $	349
Males	3
Females	2

Table 1: Statistics about Aphasia videos.  $d$ : duration,  $w$ : word

The second source consists of the transcription of the reading of a text of 131 words by Guy de Maupassant<sup>2</sup> by people with aphasia. This kind of data should be handled with care: reading difficulties might be a by-product of another language disorder frequently accompanying aphasia: alexia. More of it in the next section.

The third source is based on the transcription of two conversations between a PwA and a speech therapist (Colin et al., 2016). This allows the PwA to speak and express themselves without being interrupted.

### 3.1 Analysis of the collected data

We analyzed the transcripts to characterize the effects of aphasia on speech. Several interesting details were observed, among them we can mention that aphasia leads to hesitations, the repetition of the same word or the same syllable, the interruption of speech and the use of periphrases.

In this article, we focus on Aphasia lexical errors. Our objective is to use minimal complexity and confusability in our data as what has been done for the images of PNT (Mirman et al., 2010) in order to facilitate the rehabilitation. In lexical errors, a word form is disturbed at several levels. It may concern the replacement of a character by another (*abricot* becomes *apricot*), swapping of syllables (*télévision* becomes *létévission*). Sometimes the PwA replaced a whole word by another one. This replacement can be explained by the pronunciation proximity (*cigarette* becomes *ciguerapette*) or by a semantic confusion. For example, a word could be replaced by another one semantically close for example *pain* (*bread*) is replaced by *vin* (*wine*) and in addition, in this case these two words are acoustically close to

<sup>2</sup>Pierrot in Contes de la bécasse, 1883

each other. Sometimes the PwAs create new words, it would seem from our study that they maintain the morphology of the words.

It's worth mentioning the influence of a potential aphasia-byproduct language disorder called alexia (Cherney, 2004). There are two main types of alexia: one influences vision in a physical way, the other one damages linguistic processing. Some of the results from reading tasks might be explained by the psycholinguistic deficits caused by a degradation in linguistic processing, and are not necessarily aphasia-related.

We identified from the transcriptions of 43 erroneous words belonging to the class of lexical errors, four categories: substitution, addition, deletion and replacement errors. Figure 1 illustrates the distribution of aphasic errors according to the Levenshtein distance between the correct word and its erroneous aphasic. This figure shows that 67% have a Levenshtein distance smaller or equal to 2 with the correct word.

Therefore, based on these figures, we believe that it is possible to create a large enough aphasia corpus by simulating errors close to those we encountered when analyzing real aphasia data. This will be done by introducing type errors: insertion, deletion and substitution, based on appropriate values of the Levenshtein distance.

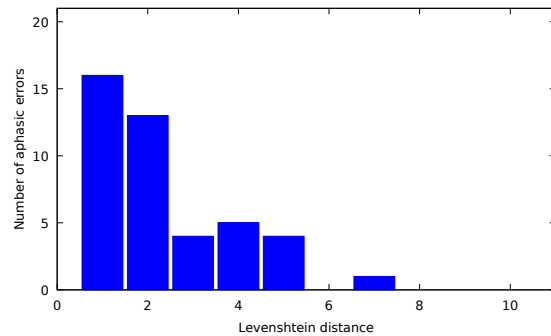


Figure 1: Distribution of aphasic errors according to the Levenshtein distance.

### 3.2 Automatic generation of pseudo-aphasic corpus APHAFRECH

Thanks to the errors studied during the analysis of aphasic data, we propose to create a pseudo-aphasic corpus automatically. In the following, we refer to all aphasic errors and their corrections, described in Section 3.1, as  $\mathcal{A}$  ( $\mathcal{A}$  for Aphasic). This corpus is made up of couples  $(a_i, c_i)$  where  $a_i$  is the  $i^{th}$

aphasic error and  $c_i$  is its correct version proposed by a human annotator.

In order to build APHAFRECH, we propose to start with  $\mathcal{C}$ , a clean corpus totally independent from the aphasia corpus described in Section 3.1. To build  $\mathcal{C}$ , we extracted 2,000 short sentences from the French part of the English-French file produced by Tatoeba project<sup>3</sup>. Then we generated from each sentence of  $\mathcal{C}$  a pseudo-aphasic sentence by applying rules based on the analysis of  $\mathcal{A}$ . For that, we apply the method described in Algorithms 1 and 2. For each sentence in  $\mathcal{C}$ , we randomly select words to alter with a fixed probability  $p$ . Then for each selected  $w$ , we produce  $n$  erroneous words potentially considered as words pronounced by a PwA. These  $n$  words are produced by using substitution, deletion and insertion of letters within  $w$ . Then, among these produced erroneous words, we select the best one  $w'$  which will replace  $w$ . In the alter function (Algorithm 2), for this first experiment, we allow only one alteration, but the algorithm can be later extended to lead to several alterations. We define this best erroneous word by using a scoring that yields the most likely altered word of having been pronounced by a PwA. With this two-step process, we want to give as much freedom as possible to the generation of errors, even if it means generating errors that are actually impossible to pronounce; step 2 then allows us to select plausible errors.

#### Algorithm 1 Generation of APHAFRECH

**Require:** a corpus  $\mathcal{C}$ ,  $p$ ,  $n$   
**Ensure:** parallel corpus APHAFRECH  
 APHAFRECH  $\leftarrow \emptyset$   
**for each** sentence  $s \in \mathcal{C}$  **do**  
    $s' \leftarrow$  empty string  
   **for each** string  $w \in s$  **do**  
   **if** random()  $< p$  **then**  
     $w' \leftarrow \text{alter}(w, n)$   
   **else**  
     $w' \leftarrow w$   
   **end if**  
   add  $w'$  to  $s'$   
**end for**  
 add couple  $(s, s')$  to APHAFRECH  
**end for**

#### Algorithm 2 Generation word errors (function alter)

**Require:**  $w$ ,  $n$   
**Ensure:** erroneous variant  $w'$  of  $w$   
 $v \leftarrow \emptyset$  (the set of variants)  
**for each**  $i$  from 1 to  $n$  **do**  
   **repeat**  
    $w_i \leftarrow w$   
   alteration  $\leftarrow$  random("I", "D", "S")  
   **if** alteration = "I" **then**  
    insert randomly a character in  $w_i$   
   **else if** alteration = "D" **then**  
    delete randomly a character from  $w_i$   
   **else**  
    replace at random position a character of  $w_i$  by another one randomly selected  
   **end if**  
   **until**  $w_i \notin v$  or a maximum number of iterations is reached  
   add  $w_i$  to  $v$   
**end for**  
**for each**  $w_i \in v$  **do**  
   give a score to  $w_i$   
**end for**  
 $w'$  is the  $w_i$  with the best score

#### 3.2.1 Scoring a variant

Each erroneous variant is assigned a score that indicates to what extent it could have actually been produced by a PwA. Then, in the initial sentence, we replaced the concerned words by those that achieve the best error scores. To measure the quality of a variant, we tested two scoring functions. Actually, the variant  $w'$  is supposed to be pronounced by a PwA and as it is difficult to affirm that, the PwA wanted to say  $w$ , we should score this word. We define a measure  $f(w, w')$  that gives values between 0 ( $w'$  is certainly not an aphasic word spoken in place of  $w$ ) and 1 ( $w'$  could be certainly an aphasic word spoken in place of  $w$ ). In the following, we define the two score measures.

**ngram scoring** For this scoring, the quality of the erroneous string  $w'$  depends only on the likelihood of the character sequence. This likelihood is computed based on a character ngram language model that has been trained on a French novel (Germinal, by Émile Zola). For the smoothing method, we used Katz method (Katz, 1987). In sake of future coverage, the character vocabulary is the set of unicode ids. We propose to define the ngram

<sup>3</sup><https://www.manythings.org/anki/>



score by Equation 1.

$$ngram(w, w') = \frac{1}{m} \sum_{i=1}^m P(w'_i | h'_i) \quad (1)$$

Where  $m$  is the number of characters of  $w'$  and  $h'_i$  is the character sequence preceding  $w'_i$ . In case of  $n$ -gram,  $h'_i$  is truncated to the  $n - 1$  preceding characters.

Let's remark that  $ngram(w, w')$  does not depend on  $w$  because we want only measure the likelihood of  $w'$  independently of the lexical distance between  $w$  and  $w'$  (this distance is fixed to 1 in this experiment).

**soundex scoring** For soundex, the words  $w$  and  $w'$  are close if their respective pronunciations are close. To estimate the degree of closeness of words, we compared the soundex encoding of  $w$  and  $w'$ . Soundex (Jacobs, 1982), is a method for indexing words by their sound. Words are encoded by taking advantage of their phonetic form. The encoding is done in both words, the altered and the correct one. The principle of encoding a word consists in deleting spaces, uppercasing the word, keeping the first letter, deleting the vowels, associating digits to each letter in accordance to its phonetic class (see Table 2) and finally by keeping the first four characters.

Phonological group	digit
B,P	1
C,K,Q	2
D,T	3
L	4
M,N	5
R	6
G,J	7
X,Z,S	8
F,V	9

Table 2: Soundex codes for each phonological group

With this encoding function, words like *ballon* and *ballon* will receive the same code B445, while the encoding of the words *farapluie* and *parapluie* will be respectively F614 and P614.

Then we estimate the Soundex closeness of  $w$  and  $w'$  by  $soundex(w, w')$ , defined by Equation 2.

$$soundex(w, w') = \frac{1}{4} \sum_{i=1}^4 \delta(S_i(w), S_i(w')) \quad (2)$$

Where  $S_i(x)$  is the  $i^{th}$  soundex code of the word  $x$ .  $\delta(x, y)$  returns 1 if  $x$  is equal to  $y$ .

**Performance of scoring functions** In order to measure the capability of ngram and soundex to give a score close to 1 to real aphasic errors, we use  $\mathcal{A}$  as a test corpus. We injected each real aphasic error  $a_i$  into the list of pseudo aphasic errors provided by Algorithm 2. Table 3 shows the average of the inverse rank of  $a_i$  in the list sorted according to the ngram and soundex scores. For ngram, we tested several values of  $n$ , the best results have been achieved for  $n = 4$ . The result of the soundex function leads to a very low performance compared to the ngram function. This is due to the distribution of the soundex function scores which has a tiny standard deviation.

Scoring function	Performance
4-gram	0.26
Soundex	0.02

Table 3: Performance of scoring functions on the produced errors

#### 4 A preliminary experience in Machine translation of a pseudo aphasic corpus

In this section, we study the opportunity to translate an aphasic corpus to its corrected counterpart. For that we use APHAFRESH, the parallel corpus we described in Section 3. To generate the aphasic sentences, we used only the ngram scoring function since it is the one that achieves the best aphasia errors. Table 4 shows a sample of this corpus.

Pseudo-aphasia sentences	Correct sentences
sois juite	sois juste
j'ai fait sine	j'ai fait signe
je dui calme	je suis calme
je me suis révemillé	je me suis réveillé
je suis detite	je suis petite
si ça ne vous dérande pas, pourrions-nous inspecter votre valise	si ça ne vous dérange pas, pourrions-nous inspecter votre valise ?

Table 4: A sample of the parallel experimental pseudo-aphasic corpus APHAFRESH

Our ultimate objective is to provide an aphasic speech to natural speech translation system. But, in this preliminary experience, we will study the

opportunity to translate an aphasic corpus to its corrected counterpart. For that, we will train a sequence-to-sequence machine translation model, a kind which has been used widely in the literature of machine translation and other NLP applications (Sutskever et al., 2014; Zhang et al., 2015; Nguyen Le et al., 2017; Mao et al., 2020). We used the corpus we created, APHAFRESH, for training, tuning and for testing.

The input of the encoder is the Aphasia sentence and the output is the hidden state and cell state of the LSTM. The decoder has the hidden state and cell state of the encoder as inputs in addition to the input sentence. The results of the decoder LSTM is passed through a dense layer to predict decoder outputs as shown in Figure 2.

In Table 5, we give the different parameters of the neural network architecture we used.

Parameters	Values
Source Maximum Length sentence	13
Target Maximum Length sentence	14
Source Unique words	13,085
Target unique words	8,364
Batch Size	64
Epochs	20
Number of LSTM Nodes	400
Embedding Size	100
SPLIT Training-Tuning	0.1
Test size	2,000

Table 5: The parameters of the sequence-to-sequence model

Concerning the optimizer, in our experiments we tested several methods, the one which achieves the better results is the Adaptive Gradient Algorithm (Duchi et al., 2011). In fact, adaptive gradient algorithms calculate gradient-based updates using the history of gradients, which has the advantage to reduce the inconvenience of manually setting the step size parameter in the stochastic gradient descent optimizer. In addition, AdaGrad is known also for its computational efficiency (Kingma and Ba, 2014). From Figures 3 and 4, we can conclude that the accuracy is high and the model reduced the value of the loss, which means that the model makes small errors on few data and the model predicts well. The training and the validation curves start with relatively high loss at the beginning and gradually decrease as training and validation examples are added and gradually flatten, indicating that

adding more examples does not improve the performance of the model on both data. This leads us to assume that our neural network does not overfit.

We tested this model in a test corpus composed of 2,000 aphasia sentences. The results in terms of cumulative BLEU are given in Table 6.

BLEU1	BLEU2	BLEU3	BLEU4
59.24	51.20	44.39	38.60

Table 6: Cumulative BLEU on the pseudo-aphasia corpus

Figure 5 illustrates the distribution of the BLEU score over the 2,000 sentences of the test corpus. We can observe that more than 31% of the sentences have a BLEU higher than 50 which means that we achieve a very high quality and fluent translation. Only 5% of the translation have a BLEU smaller than 10 which corresponds in general to a useless translation. 19% of the translations have a BLEU between 10 and 19, which corresponds to sentences that are difficult to understand.

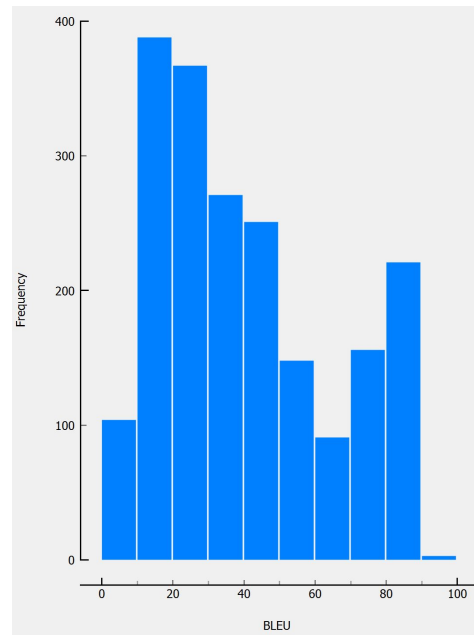


Figure 5: Distribution of the level of understanding of the translation of the 2000 Aphasia sentences

In order to make the reading of Figure 5 easy, Table 8 recalls how to interpret the BLEU score (Noever et al., 2021) accordingly to the quality of the translation.

The global analysis of the BLEU score on the different sentences of the test corpus is illustrated by Table 7.

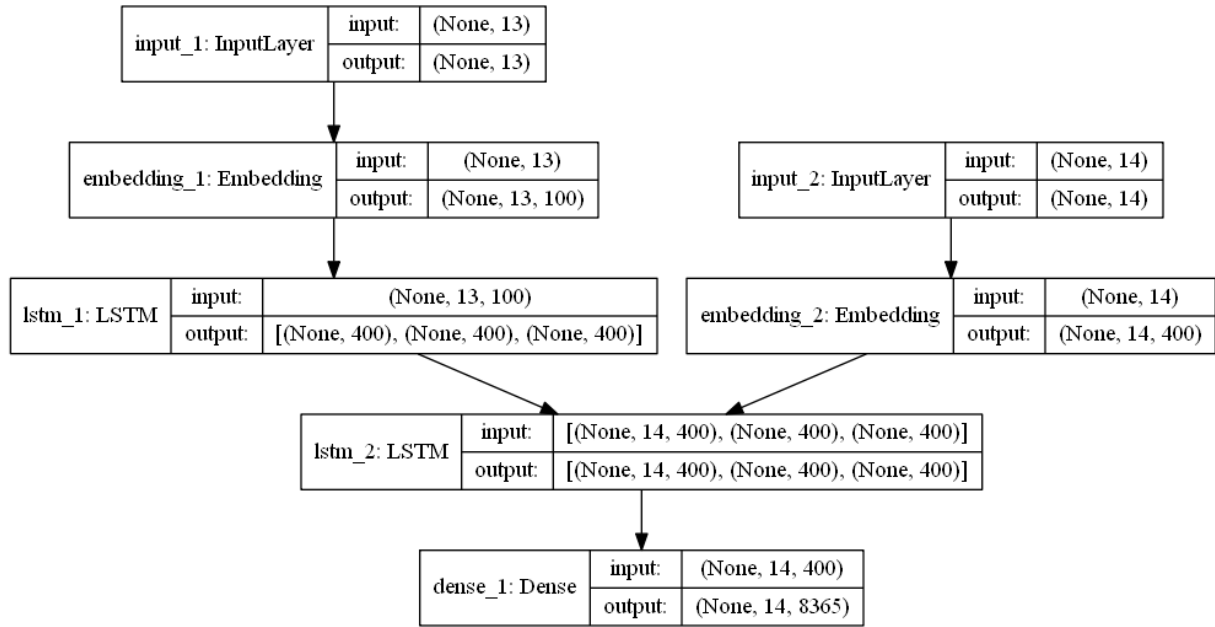


Figure 2: Architecture of the Aphasia sequence-to-sequence model

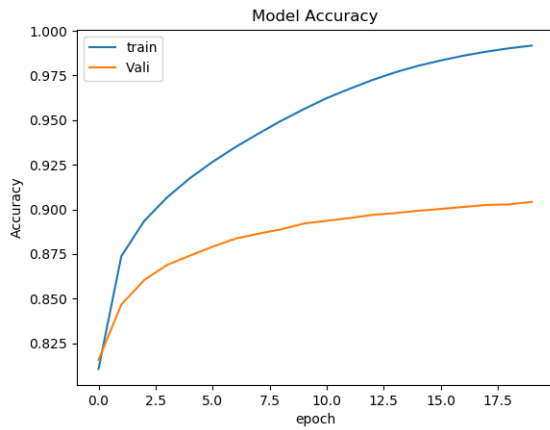


Figure 3: The accuracy on the training and the validation corpus

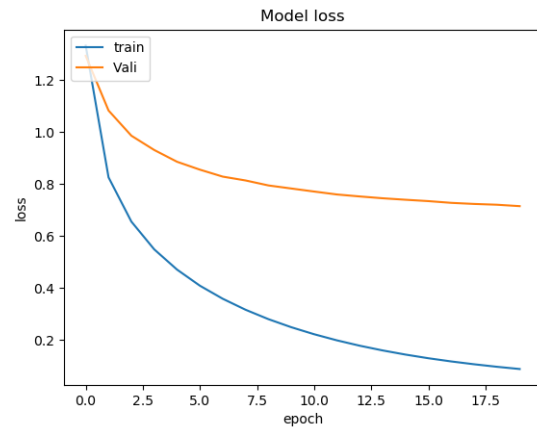


Figure 4: The loss function on the training and the validation corpus

Mean	SD	Max	Min
40.62	24.09	90.48	4.18

Table 7: Some figures concerning the BLEU scores of the Aphasia to natural text machine translation

Figure 6 is a different presentation of Figure 5, it shows the decreasing evolution of the values of BLEU. We can notice that more than 25% of the test corpus was translated with a performance of at least 50 in terms of BLEU.

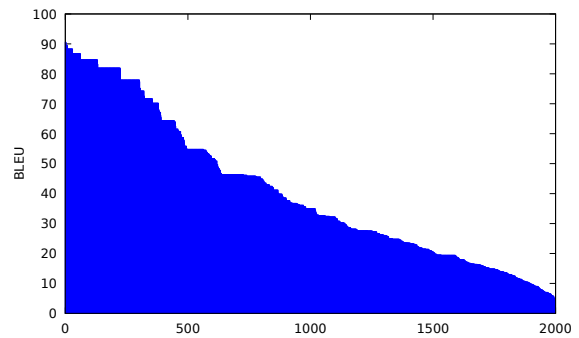


Figure 6: A decreasing distribution of BLEU over 2,000 sentences.



BLEU	Score Interpretation
< 10	Almost useless translations
10 to 19	Difficult to understand
20 to 29	The idea is clear, but it contains many errors
30 to 40	correct translations
40 to 50	High quality translations
50 to 60	Very high quality
> 60	Quality better than a human translation

Table 8: Interpretation of the quality of machine translation depending on the BLEU score

## 5 Conclusion

Aphasia is a unique and rather complex phenomena. There is a great amount of work trying to understand and explain the underlying structural changes from different perspectives. Since there is no general consensus on what the best approach is to therapy, the field remains open for experimentation. We decided to take up the challenge from a machine learning perspective by implementing a method that will eventually allow us to come up with a speech-to-speech system where the input is aphasic speech and the output is a rehabilitated speech. For that, we created an aphasia-like corpus, APHAFRECH, with correct-incorrect sentence pairs, using three different resources. This required the study of errors from aphasic sources in order to understand certain types of errors and to reproduce them automatically. With the created dataset we trained a neural network machine translation that yields very high quality translations on APHAFRECH. The next step will concern the introduction of more complex aphasia errors into APHAFRECH (such as context-dependant errors and semantic based errors) and the study of the quality of the translation by using a more elaborated DNN machine translation.

## References

- Elizabeth Armstrong. 2000. [Aphasic discourse analysis: The story so far](#). *Aphasiology*, 14(9):875–892.
- Leora Reiff Cherney. 2004. Aphasia, alexia, and oral reading. *Topics in Stroke Rehabilitation*, 11(1):22–36.
- Soojin Cho-Reyes and Cynthia K. Thompson. 2012. Verb and sentence production and comprehension in aphasia: Northwestern assessment of verbs and sentences (navs). *Aphasiology*, 26(10):1250–1277.
- Sharice Clough and Jean K. Gordon. 2020. Fluent or nonfluent? part a. underlying contributors to categorical classifications of fluency in aphasia. *Aphasiology*, 34(5):515–539.
- Capucine Colin et al. 2016. Adaptation du projet aphasiabank à la langue française: Contribution pour une évaluation informatisée du discours oral de patients aphasiques.
- Marjory Day, Rupam Kumar Dey, Matthew Baucum, Eun Jin Paek, Hyejin Park, and Anahita Khojandi. 2021. Predicting severity in people with aphasia: A natural language processing and machine learning approach. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pages 2299–2302.
- John C. Duchi, Elad Hazan, and Yoram Singer. 2011. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.*, 12:2121–2159.
- Margaret Forbes, Davida Fromm, and Brian Macwhinney. 2012. AphasiaBank: a resource for clinicians.
- Maria Garraffa and Valantis Fyndanis. 2020. [Linguistic theory and aphasia: an overview](#). *Aphasiology*, 34(8):905–926.
- J.R. Jacobs. 1982. Finding words that sound alike. the soundex algorithm. *Byte* 7, pages 473–474.
- Slava Katz. 1987. Estimation of probabilities from sparse data for the language model component of a speech recognizer. *IEEE transactions on acoustics, speech, and signal processing*, 35(3):400–401.
- Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization.
- Zhuoyuan Mao, Fabien Cromieres, Raj Dabre, Haiyue Song, and Sadao Kurohashi. 2020. Jass: Japanese-specific sequence to sequence pre-training for neural machine translation. *arXiv preprint arXiv:2005.03361*.
- Randolph S Marshall, Ronald M Lazar, and JP Mohr. 1998. Aphasia. *Medical Update for Psychiatrists*, 3(5):132–138.
- Daniel Mirman, Ted J. Strauss, Adelyn Brecher, Grant M. Walker, Paula Sobel, Gary S. Dell, and Myrna F. Schwartz. 2010. A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function. *Cognitive Neuropsychology*, 27(6):495–504.
- An Nguyen Le, Ander Martinez, Akifumi Yoshimoto, and Yuji Matsumoto. 2017. Improving sequence to sequence neural machine translation by utilizing syntactic dependency information. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 21–29, Taipei, Taiwan. Asian Federation of Natural Language Processing.

D. Noever, Josh Kalin, Matthew Ciolino, Dom Hambrick, and Gerry Dozier. 2021. Local translation services for neglected languages. pages 149–163.

Elizabeth Rochon, Laura Laird, Arpita Bose, and Joanne Scofield. 2005. [Mapping therapy for sentence production impairments in nonfluent aphasia](#). *Neuropsychological rehabilitation*, 15:1–36.

Debbie Summers, Anne Leonard, Deidre Wentworth, Jeffrey L Saver, Jo Simpson, Judith A Spilker, Nanette Hock, Elaine Miller, and Pamela H Mitchell. 2009. Comprehensive overview of nursing and interdisciplinary care of the acute ischemic stroke patient: a scientific statement from the american heart association. *Stroke*, 40(8):2911–2944.

Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc.

Charalambos Themistocleous, Kimberly Webster, Alexandros Afthinos, and Kyrana Tsapkini. 2021. Part of speech production in patients with primary progressive aphasia: An analysis based on natural language processing. *American journal of speech-language pathology*, 30(1S):466–480.

C. K. Thompson, K. J. Ballard, M. E. Tait, S. Weintraub, and M. Mesulam. 1997. [Patterns of language decline in non-fluent primary progressive aphasia](#). *Aphasiology*, 11(4-5):297–321.

Cynthia Thompson, Lewis Shapiro, Swathi Kiran, and Jana Sobecks. 2003. [The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia](#). *Journal of speech, language, and hearing research : JSLHR*, 46:591–607.

L. Williams, M. Ali, K. VandenBerg, J. Godwin, A. Elders, F. Becker, A. Bowen, C. Breitenstein, M. Gandolfi, E. Godecke, K. Hilari, J. Hinckley, S. Horton, D. Howard, L.M.T. Jesus, M. Jungblut, M. Kambanaros, T. Kukkonen, A. Laska, B. MacWhinney, I. Martins, F. Mattioli, M. Meinzer, R. Palmer, B. Patrício, C. Price, N. Smania, J. Szaflarski, S. Thomas, E. Visch-Brink, L. Worrall, and M. C. Brady. 2016. Creating an international, multidisciplinary, aphasia dataset of individual patient data (ipd) for the rehabilitation and recovery of people with aphasia after stroke (release) project. *International Journal of Stroke*, 11(4).

Jiajun Zhang, Chengqing Zong, et al. 2015. Deep neural networks in machine translation: An overview. *IEEE Intell. Syst.*, 30(5):16–25.