IPRPD
International Journal of Arts, Humanities and Social Sciences
ISSN 2693-2547 (Print), 2693-2555 (Online)
Volume 07; Special Issue: March 2026
DOI: 10.56734/ijahss.v7ns4a7

GENERATIVE AI FOR WRITING IN FRENCH OR ENGLISH

Claire Mathieu[1]

[1] CNRS, Université Paris Cité

Abstract

How can we understand the spectacular successes and strange failures of the tools of generative artificial intelligence for instantaneous translations of texts between two languages ?

First I will present the algorithmic methods underlying large language models (LLMs) such as ChatGPT, DeepL or Claude, and then I will describe a recent experience that I had using them for translation..

Keywords

Large Language Models, Translation, Algorithm, Neural network

How can we understand the spectacular successes and strange failures of the tools of generative artificial intelligence for instantaneous translations of texts between two languages ?

First I will present the algorithmic methods underlying large language models (LLMs) such as ChatGPT, DeepL or Claude, and then I will describe a recent experience that I had using them for translation.

The methods underlying large language models

What is the method underlying Large Language Models? During a first phase, the system is trained extensively on a large corpus of documents, typically written in the English language. The goal is to generate new texts that imitate the texts in the corpus as closely as possible. What does that mean? Say that you set aside a small sample of the corpus. Then you proceed to train the LLM on the rest of the corpus. Once the training is finished, you can use the LLMs to generate new texts. The method is successful if an outside observer that sees some texts that are generated artificially and some that are taken from the sample is unable to determine which texts are generated by the machine and which are written by a human. Imitation is the only goal.

How does the training operate ? The features of the language, for example its grammatical rules, are not programmed explicitly but learned from examples. A text is decomposed into tokens, such that each word consists of one or a few tokens, possibly corresponding to the prefix, the root, and the suffix, and the tokens are generated one by one. For our purpose here, it suffices to think of tokens as words. Each is encoded by a vector, that is, a list of numbers, but the encoding is not fixed. We can think of the encoding as a kind of mental image associated to the word. For example, with no information about the context, the word “dog" might be represented by a list of numbers such as (.5,.7,.6), but if it occurs in the sentence “After playing fetch, the tired dog is resting in the shade”, then the word “dog" in that context might be represented by the list (.4,.8,.7); a different list, because of the context of nearby words that give information adding some nuance to the meaning of “dog”. This is not just any dog, but a dog who has fetched and who is resting. If we continue the text, in the following sentences we might learn that the dog is panting, and the context makes the word “panting" more likely to come up. Since that is true in the corpus, if the training is successful then that will also hold for the text generated artificially by the LLM.

The reader may have observed that when we use a LLM, asking the same question twice may lead to different answers. How can that be? The reason is that the text generation is probabilistic. The underlying neural network is designed to generate the text word by word (or rather, token by token). Each new word is generated after the text generated so far, whose generation has led to an encoding of the story so far in the form of a list of numbers. Then, the current state is used to generate a new word according to some probabilities. For example, if the text starts with “the boy ran after the”, this text, once ingested by the LLM, leads to a certain list of numbers giving its current state, and then a new word is generated according to some probabilities: perhaps the continuation of the sentence could be “ball”, “dog”, or “cat”, and a coin is flipped by the machine to choose among those three possibilities. Once the

decision is made, that word then becomes part of the text, enriching the context and giving more information to generate the next word after that, and so on and so forth.

How does the LLM know that what it is writing is correct? The short answer is that it doesn't. In the design, there is no notion of the result being "correct". The ultimate goal of the LLMs is to generate something that resembles the documents with which it has been trained, as closely as possible. The text generated should be a plausible text of the corpus. Thus, any bias or error in the dataset, and any cultural dimension, will be naturally incorporated and reflected in the texts generated. If some part of the culture is not represented the the documents, then the same will be true in the text created. For that reason, knowing what the corpus consists of is fundamental to any evaluation of the LLM. Note that, even if the corpus had somehow been trimmed to contain only texts that are free from factual errors, that does not, by any means, insure that the texts generated will be free from error. Indeed, the LLM has no notion of truth, and no understanding of "meaning". Imagine you were looking at antique Egyptian hieroglyphs, that you noticed some patterns, and that you used them to make up your very own Egyptian text. It may well look plausible, and it might even be understandable by an Egyptian scribe, if you have spent enough time and efforts figuring out and reproducing the patterns that you see. But you have absolutely no idea what your text is about, and for you, it has no grounding in reality. It is just a meaningless sequence of symbols that resembles the ones in your corpus.

In the last century, the mathematician Alan Turing proposed a way to define artificial intelligence: imagine that a tester is holding two conversations, one with a human and the other with a computer, each hidden behind a curtain. In order for the dialogue to proceed, the tester types on a keyboard and reads the answers on the screen. After the conversation is over, the tester must decide which curtain hides the human and which one the machine. If the tester is unable to figure it out, then one can say that the machine has achieved artificial "intelligence". This is a purely functional definition. It has nothing to do with any kind of "understanding" but only with the ability, on the part of the machine, to simulate the responses of a human person. Such is the goal of the neural nets used by LLMs: somehow manage to follow, for example, the grammatical rules of the language, without being told or "understanding" anything about those rules, but purely by imitation. But what do we mean by "understanding"? One might recall the water scene in the Helen Keller movie. Helen Keller was both deaf and blind, and her teacher tried to make her learn to write words. She was mindlessly writing the words according to her teacher's instructions, when at one point, she suddenly *understood* that the word "W-A-T-E-R" was referring to water. That connection between symbols and a perception of reality does not exist in the world of large language models. The LLM has no notion of a physical object and is just reading a sequence of meaningless words strung together, and it exploits its memory and computational capacities to continue stringing words together in a plausible way.

When chatGPT came out, I tried to ask it about additions. What is the sum of 2497 and 7689, I asked? It turns out that beyond 2 or 3 digit numbers, the results were completely wrong. Why not? Because the neural network may memorize a few additions of numbers that occur frequently enough in the corpus that the result of the addition will be encoded in its representation of the world. But it will not have memorized infrequent additions of large numbers. If it were designed specifically to do additions, the neural net, as a computational model, would have the capability of adjusting some of its parameters so as to do additions mostly correctly; but the goal of the training is plausibility, not mathematical rules, so there is no *a priori* reason why some of the neurons would be reserved for arithmetic.

In the corpus on which the LLM has been trained, there are entire encyclopedias. When you ask for the kind of information that is in an encyclopedia and get a response, it is tempting to believe that the LLM "knows" the answers to the questions you are asking, but that impression is misleading. There is not enough time and space for the LLM to remember everything perfectly and retrieve it instantaneously. If it memorized every single thing it had been trained on, it would run out of memory and of computing power to find the answer. So, instead, it encodes its knowledge with number lists, that provide a sketch of its view of the world. Thus, in that concise view, some of the information is lost. That is why, when answering a question leads to a re-generation of the answer, some mistakes may occur, even if there are no mistakes in the corpus. Moreover, some recent theoretical work pointed out some structural limitations: a LLM that generates a faithful sample of text, not narrowed to some narrow subset but about as diverse as the texts on which it has been trained, will necessarily have a significant risk of venturing outside the world of reasonable texts. In other words, either it produces an impoverished version of language, or it risks producing some garbled, nonsensical texts. One cannot have faithful coverage of existing content without risking to generate implausible texts. Thus, efforts to limit the number of errors necessarily lead to texts that are less rich and diverse than human-generated documents.

Translating English into French with LLMs

Now that we have discussed some of the principles underlying the design of large language models, let us examine one application, to automated translations. That is a basic application of LLMs, to which they are particularly well-suited. I would like to report on my experience with that.

LLMs were first trained on the English-language corpus. How did they also learn French? The designers did not start all over with a French language corpus. Once extensive work has been done to train the LLM in English, it is not necessary to repeat everything from scratch in order to create a generative AI that can produce texts in French. With a dictionary that gives some basic correspondences, plus some texts that already exist in bilingual editions and provide the specific French context for some famous French sayings, that already provides enough information for the system to be able to generate texts in French that are grammatically correct and that use native French expressions correctly. The more French texts are present in the training set, the more likely it is that the generated texts will capture some cultural nuances that are specific to French writers. It is impressive to think that once the training has been done once by working extensively in English, the additional work needed to train the LLM in a different language such as French is comparatively small. In a way, one can think of polyglots who, after learning three or four languages, are able to learn a new language with remarkable ease. They seem to be able to fit the new language in their mental map of languages, leveraging their knowledge of other languages.This is a strength of LLMs.

However, the fact that for LLMs the bulk of the documentation used for basic original training is in English shows in ways that are more subtle than merely awkward sentence constructions.

For example, think about what secondary school students are taught about writing. The basic method taught in high school in France for writing dissertations follows the three part plan “thesis-antithesis-synthesis”, and that education shows its influence well beyond the school context, permeating the texts written by native French people. In contrast to that approach, in the US, in middle school one generic method teaches students to write texts consisting of five paragraphs. Each paragraph contains about five sentences: it first presents an idea, then might illustrate it by some example, and finally ends with a sentence summarizing the content of the paragraph. Again, this method influences the writing of many texts in the English language. The reader who is familiar with both cultures can then read a text in English, and, even without the knowledge that the text is a translation, and even if the translation is flawless, may be able to recognize from the underlying structure that is the signature of a French “way of thinking”. Because LLMs are first trained in English, they are permeated with an English “way of thinking”. Even if the texts that are generated in French are spotless in their wording, they still unmistakably come from a different culture.

More concretely, I was recently exposed to an attempt to translate a Mathematical textbook from English to French using LLMs. I had been asked to review the translation of the first chapter. The results turned out to be worse than what I had expected, in several different ways.

First, the (generic) LLMs were unable to deal with mathematical notation, and they garbled the equations. Perhaps that was to be expected since they were general-purpose and had not had a training focused on those. An expression such as “f(x)=y” is perhaps frequent enough in documents that it can be accepted, but the opening parenthesis is not preceded by a space, and that throws off the LLM by moving it to a region of its representation space where typesetting rules do not apply, which may lead to strange things happening afterwards as it continues generating text.

Secondly, in the textbook there was a joke about even and odd numbers, based on the two meanings of the word odd: “odd” as in “not divisible by two”, and “odd” as in “strange”. The LLMs missed the joke completely and the corresponding text in French made no sense. Actually, I would have been impressed if the LLMs had “understood" the joke, since its representation (“understanding”) of words comes from context, and in the context of that chapter, the word “odd" always referred to numbers no divisible by two, with the exception of one single occurrence for the joke. That unexpected change in meaning is precisely the root of the joke. Given the method used by LLMs, those types of jokes seem ideally suited to fooling them. A human translator would perhaps have created a different joke based on the two meanings of the word “impair" in French: “impair” as in “not divisible by two”, and “impair" as in “awkward mistake”. But that would require a much higher level of reasoning: not just blindly imitating existing texts, but recognizing the presence of the two meanings of “odd”, realizing that it creates a humorous effect, that the humorous effect is the main goal of the writer in that paragraph, searching for a mathematical word frequently used in the chapter and whose French translations has two meanings, and imagining a short French text that uses both meanings of that word.

Thirdly, the chapter contained some specialized mathematical terminology, in the form of words that are used in everyday language but that, in Mathematics, have a very precise meaning; and the LLMs translated several of those words by the equivalent French everyday word instead of the unique correct translation, namely, the mathematical word that in French is used to characterize the mathematical structure described. The textbook was meant to be for teenagers, addressing them in informal ways, with engaging stories along the way, so it mixed some ordinary language with some precise mathematical wording. To a human translator, it would have been obvious which words were referring to a mathematical concept and therefore needed an exact translation into the proper French word for that concept, but the LLMs were unable to tease out the mathematical registry from the ordinary language registry.

Fourthly, there are some mathematical words that are misleading (“false friends”, as we say in French). Sometimes, there are mathematical notions that can be defined in several different ways, and the choice is arbitrary; it can happen that mathematicians from different countries made different choices. One example is the word “positive" and its French analog “positif”. In English, a positive integer is an integer strictly greater than zero: 1,2,3,

etc. In French, “entier positif” refers to an integer greater than or equal to zero: 0,1,2,3, etc. Close, but different! The LLMs blindly translated “positive" into “positif”, and as a result, some of the proofs became incorrect. For a mathematical textbook, that outcome is an outright disaster. That is perhaps the most forgivable of the mistakes I saw, because many non-scientific human translators would have made the same mistake. To avoid that kind of fatal translation error, a human translator needs to know enough Mathematics to either be aware of the subtle distinctions between the definitions of the words in the two languages, or to be a French translator who is comfortable enough with Mathematics to understand the proof that they are translating and realize that if they translate the word “positive" by “positif”, then the resulting French-language proof is wrong.

Altogether, those examples illustrate the weaknesses of a language model that is based on imitation rather than understanding. They also point to one main weakness of material generated in French using a LLM: since the training of the LLM was done with English-language documents, the LLM is imitating French text, but within a framework broadly educated by an English culture and perspective, leading to a different way to organize one’s thoughts, present arguments, and ultimately a different way to think. When one reads a text written by someone who has lived in France for many years and who speaks perfect French, but who was educated in the US, sometimes one thinks: “This is a very American way to phrase things”. The LLMs have been primarily educated with English-language media, and even if they learn to speak flawless French in a perfect imitation of French speakers, they are still more broadly modeled after English-language texts and literature. That means that the increased use of LLM-generated texts will lead to a loss of cultural diversity. The LLM invaders, on the surface, have put on the superficial garments of French words and expressions, but the background that is part of the French identity and that shapes our way of thinking and of organizing our thoughts will gradually fade, in ways that are almost imperceptible. What used to be called “*le génie français*” might soon be in danger.

What does this tell us about media? I conjecture that the same holds for videos. For example, I expect that characters generated by AI will smile with open lips, showing their teeth, in the American way, instead of smiling with their lips closed, as the French are prone to doing. I expect that they will stand next to one another at a distance that is comfortable in the American culture, but different from what it would be in some other culture. It does not mean that there is anything wrong with smiling with open or with closed lips, but that in subtle ways, in small and large things, without being fully aware of it, we will be fed the underlying American culture of videos that the AI systems have been primarily trained on.

To summarize, LLMs have been aptly described as stochastic parrots. They imitate the corpus on which they have been trained without understanding. That is not meant as a put-down, because sophisticated imitation is more powerful than we could ever have expected. Unfortunately, it requires a lot of resources, and leveraging the work done in English to write texts in French means that the British and US cultures will indirectly infuse those French-looking texts. We need to develop an awareness of those influences, so that we may preserve the diversity of our human cultures.