Cambridge University Press & Assessment
978-1-009-55236-3 — Lost in Automatic Translation
Vered Shwartz
Excerpt
More Information

Introduction

Picture, for a moment, enlisting the help of automatic translation when you seek medical attention in a foreign country and need to explain, in no uncertain terms, where you experience pain and in what intensity. I have experienced this in my first year in the US after moving there from Israel. Now consider that I'm not only a user of language technologies but also a researcher working on these technologies. As such, I'm also aware of their limitations. For example, I know that translation systems may translate figurative expressions literally, or that certain inputs can make them generate incorrect "translations" in the form of a religious text.

English is not only the world's foremost go-to language for communication and collaboration on research, information, and ideas. It also dominates the internet, which started as a network for researchers in the US. And given that the most dominant tech companies are US-based, language technologies – including automatic translation and personal assistants – tend to be English-centric.

Mastering English past the point of literal translation has been as crucial for me in my work as a natural language processing expert and for my life in English-speaking countries – first in the US and then in Canada. Many of my fellow English as a foreign language (EFL) colleagues feel the same. We've all embarked on a journey that included acquiring vocabulary and learning to form grammatical sentences as well as using and understanding figurative expressions, euphemisms, cultural references and norms, and nonverbal communication. Despite achieving a level of proficiency where we can confidently articulate our thoughts in English, our accents may still give us away. It can be frustrating when we cannot make ourselves fully understood, when a person can't make out the words we speak, or when Siri or Alexa fail to activate until we fake an American accent (one of the many occasions where my accent lessons paid dividends).

The goal of passing as a native English speaker can be challenging to reach, and I have collected numerous examples of the frustrating but often humorous experiences from the EFL community navigating life in an English-speaking world. Beyond the entertainment value of recounting such anecdotes, they reveal deeper insights that can help advance our understanding of linguistics, natural language processing, and language education.

As a means for conveying and preserving shared values and traditions, language is intrinsic to the expression of culture. By extension, language learning can help foster empathy and bridge differences – an important function in today's globalized society. What then is the role of language technologies, such as personal assistants like Siri and Alexa, automatic translation, and chatbots? Will they aid us in the quest to master foreign languages and better understand one another? Or will they make language learning obsolete?

The last decade has seen an exponential increase in the development and adoption of language technologies, including generative artificial intelligence tools like ChatGPT and Gemini. While such technologies appear capable of answering sophisticated questions with high accuracy and even generating creative texts like poems, their answers, at times, can be incorrect or inconsistent despite sounding confident and plausible. What's more, the technology doesn't really "understand" language the way humans do (and it is certainly not sentient).

As a natural language processing expert, I believe language technologies, in general, can help improve communication capabilities for both native and non-native speakers. Yet questions remain about how reliable they are, what impact they will have – for example, on the labor market, education, and society as a whole – and what we stand to lose or gain when we count on them without addressing their issues.

This book is the result of my deep love and respect for the diversity of human language and its power to enable us to learn more about each other, understand different perspectives, and work together to build a more inclusive global community. It's my hope that through a thoughtful approach, and a healthy sense of humor, language technologies can help rather than hinder us in this quest.

Cambridge University Press & Assessment
978-1-009-55236-3 — Lost in Automatic Translation
Vered Shwartz
Excerpt
More Information

PART I

Communicating in English

CHAPTER 1

Can We Have a Word?

London, 1971. A Hungarian man enters a tobacco shop, looking to buy cigarettes and matches. The man doesn't speak English, so he uses an English–Hungarian phrasebook. Unfortunately, for some reason, the English translations in the phrasebook are wrong. The customer tries to ask for matches, but instead says nonsensical phrases like "My hovercraft is full of eels" and "My nipples explode with delight." The frustrated salesperson tries to understand the customer, but things go south when he looks up the Hungarian translation for the price in English. When he voices what is translated into an offensive Hungarian sentence, this earns him a punch in the face from the Hungarian customer.

This is the plot of the *Dirty Hungarian Phrasebook* sketch from Monty Python, and it ends with the customer's arrest by the police. Although the sketch is grossly exaggerated – and even though physical dictionaries may be outdated – this scene remains my favorite example of the risks of communicating in a foreign language using an unreliable source of translation.

Fast forward to today, when automatic translation services such as Google Translate and DeepL have led to physical dictionaries becoming, for the most part, obsolete. Yet while such services do a pretty good job in translating between languages, it is generally considered a bad idea to rely on them unconditionally for writing a text in a foreign language you don't understand.

True, it is unlikely that innocuous sentences will be translated to completely unrelated ones, such as Monty Python's "Drop your panties, Sir William" (which is another phrase in the dirty Hungarian phrasebook). Nevertheless, evidence suggests that there are countless scenarios of automatic translation gone wrong, even some that similarly end with an arrest. In 2017, a Palestinian construction worker who was working in Israel posted to Facebook a photo of himself leaning against a bulldozer along with the text "يصبحهم." This phrase, pronounced "ysabechhum," literally

Cambridge University Press & Assessment
978-1-009-55236-3 — Lost in Automatic Translation
Vered Shwartz
Excerpt
More Information

means something like "may God bless them" and was meant as a good morning greeting. Unfortunately, Facebook's automatic translation translated it as "attack them" in Hebrew and "hurt them" in English, leading to the man's arrest by Israeli police. The source of confusion could have been the similarity to the phrase "يذبحهم," pronounced "ydbachhum," which translates to "slaughter them." As I'm writing this in 2024, I turned to the website Reverso to understand how the phrase "ysabechhum" is used in Arabic, and it was incorrectly flagged as inappropriate context – likely for the same reason.

Most translation mishaps are less disturbing – and can be funny. A Romanian relative who doesn't speak English once wished us a "happy birthday" on December 31 (coincidentally, his own birthday), because the generic Romanian greeting "La multi ani" (literally "many years") is used for wishing both "happy birthday" and "happy new year."

There are many examples on the web where automatic translation of restaurant menus results in hilarious descriptions of dishes. With a quick search, I came across a Chinese menu with an item whose description had been translated to "Fuck the duck until exploded" (please don't) and a viral Twitter post from an American visitor in a hotel in Saudi Arabia asking for help in deciphering a menu with cryptic English translations, including "She is suspicious of cheese; Not a problem; A period of cream."

Automatic translation systems have improved immensely in the last several years, but they don't perform perfectly on every pair of languages and for every type of text. Specifically, they are likely to translate full sentences more accurately than short descriptions such as menu items. While such translation fails should be reason enough to be cautious with automatic translation, let me explain how these systems work – and what their limitations are.

Machine Translation: Is It Rendering Language Learning Obsolete?

Automatic translation, also known as machine translation, started in the 1950s. During the Cold War, IBM developed a system that could translate Russian texts to English. In the early days of machine translation, each pair of languages required human translators to develop lexicons and grammar rules and programmers to code these into the software.

The next generation of automatic translation, in the early 2000s, eliminated the need to rely on human translators to develop the software. Instead, these systems relied on parallel texts in the source and target languages, such as book translations – leveraging the existing labor of human translators. Parallel text sources became the only requirement for developing a translation system for a new pair of languages, so hiring someone proficient in both languages was no longer necessary. For example, translating from Hungarian to English required what linguists call a corpus – a language resource consisting of a large and unstructured set of texts – in Hungarian and its translations in English.

A basic algorithm, which can be applied to any pair of languages, would go through Hungarian sentences and their human-translated English equivalents. For a given pair of sentences, the algorithm then aligns Hungarian phrases with their English counterparts. It counts how many times each phrase in Hungarian is translated into each English phrase throughout the entire text. A phrase in Hungarian might have several translation options to English, and each option is scored according to the number of times the phrases appear in parallel.

Differences in translation of certain phrases can be the result of ambiguity in the source language. For example, the Hungarian word *kormány* means both "government" and "steering wheel" depending on the context. They can also happen due to *lexical variability* in the target language; for instance, the Hungarian word *különböző* might be translated in English to any of the synonyms of "different," such as "diverse" or "distinct."

With each sentence in Hungarian, the system would go through the phrase translation table and come up with multiple English sentence options based on the various English translations of each Hungarian phrase. It would then choose the best translation according to two criteria: faithfulness and fluency.

First, the translation should be as faithful as possible to the original Hungarian sentence. This may be achieved by translating each Hungarian phrase to what is designated as the most frequently used corresponding English phrase in the training corpus. However, this may result in an ungrammatical or nonsensical English sentence. For example, if "government" is a more frequent translation of *kormány* than "steering wheel," the system might incorrectly translate the Hungarian equivalent of "Where is the steering wheel in this car?" to "Where is the government in this car?" in English.

To balance this criterion, the system also optimizes fluency in the target language. Fluency is measured with an English *language model* that estimates the probability of producing a given sentence in English. A simple and familiar illustration of a language model is auto-complete on your phone. You type a sentence in English and the phone suggests the most likely next word. A language model may be used to compute the

Cambridge University Press & Assessment
978-1-009-55236-3 — Lost in Automatic Translation
Vered Shwartz
Excerpt
More Information

probability of a sentence in English by computing the product of probabilities of each phrase given the beginning of the sentence. Language models capture interesting language phenomena. At the very basic level, a grammatically correct sentence such as "he eats pizza" would yield a higher score than the grammatically incorrect sentence "he eat pizza." Language models even capture some logic, such as scoring "it's raining outside and the ground is wet" higher than "it's raining outside and the ground is dry," and cultural norms, such as scoring "good Italian food" higher than "good British food."

Another major advancement in automatic translation happened in 2016 with the switch to neural network-based methods. An artificial neural network, which we will refer to throughout the rest of the book simply as a "neural network," is a method in artificial intelligence that learns from examples to recognize patterns in the data and make predictions accordingly. This approach is inspired by real neural networks in the human brain, which consist of connected neurons.

An artificial neural network gets an input, which is represented as an array of numbers. The input goes through layers of interconnected (artificial) neurons that transform these numbers by multiplying them with the "weight," a number associated with each neuron. If the output of any individual node is above the specified "bias," another number which serves as the threshold value for each neuron, the node sends data to the next layer of the network. The last layer is the output layer, containing the output or the prediction of the network.

To give a concrete example, one can train a neural network to recognize whether a certain email is spam or not. The input in this case would be an array of numbers representing the email. For example, we can count how many times each word in the English language appeared in the email. You can guess that certain words such as "win" and "free" would tend to appear more in spam emails, so predicting whether an email is spam or not based on the words inside the email is a reasonable thing to do. The output of this network would be a single number indicating how likely the email is to be spam.

The appealing property of neural networks is that they can learn various functions from examples – predicting if an email is spam or not, recognizing objects in an image, predicting whether a patient has a certain illness, and so on. All they require is data, in the form of inputs and their corresponding outputs. To be able to use a neural network for performing a specific task, it first needs to be trained. During training, the network observes training inputs and expected outputs and calibrates its weights and biases to correctly predict the expected outputs. Once trained to perform a specific task, the network can be applied to new inputs to predict an output.

To go back to translation, neural translation systems, like the previously described statistical translation systems, also rely on parallel text resources. However, instead of using the same algorithm for all pairs of languages, the system learns from the data a custom translation function for each pair of languages. The system's architecture is based on two neural networks. The first network, the encoder, gets a Hungarian sentence and encodes it into a vector – an array of numbers which captures the sentence's meaning but is indecipherable by people. The second network, the decoder, receives this vector that conveys the meaning of the Hungarian sentence and turns it into English, word by word. To add a new pair of languages, all the programmers need to do is train the network on a parallel body of text. With enough parallel texts to train the network, it can become an optimal translator with the ability to perform well on unseen sentences.

The release of Google Translate's neural models in 2016 led to significant performance improvements: a "60 percent reduction in translation errors on several popular language pairs" [1]. The language pairs were English to Spanish, English to French, English to Chinese, Spanish to English, French to English, and Chinese to English. All these languages are considered high-resource languages, or in simple terms, languages for which there are massive volumes of texts available, for example, from book translations and Wikipedias.

Much more challenging are low-resource languages, that is, languages for which there is not enough text available on the web. Neural networks are data-hungry and training them with a small amount of data isn't likely to result in an optimal solution.

In 2018, popular media expressed worries about Google Translate spitting out some religious nonsense, completely unrelated to the source text. At the time, I demonstrated this phenomenon for Igbo, a low-resource African language spoken primarily in southeastern Nigeria, as the source language. I am not an Igbo speaker, and I wrote what is clearly not an Igbo sentence: "i i i i i i i i [. . .]" – seventy-six i's separated by spaces. A human translator presented with such an input would respond along the lines of "I'm sorry, this is not a valid Igbo sentence." Ideally, a translation system should do the same. However, Google Translate instead presented me with the following English text: "As it is written in the book of the law of Moses, which was in the wilderness, which was before the man who did the work of the kingdom of Israel." A slightly different gibberish input in Igbo was translated into the question: "Who has been using these technologies for a long time?" Hopefully not the Igbo speakers looking to translate their words into English.

It is not a coincidence that automatic translation systems often translate phrases from low-resource languages into unrelated religious texts in English. After all, they are trained on pairs of sentences, such as a source sentence in Igbo and a target sentence in English. What they are not trained to do is recognize inputs that are not valid Igbo sentences. It would be much more useful if the translator could spot such cases – and respond with something like "I honestly have no idea what you want from me." Instead, the translator always assumes the input is valid. Even when it is given unrecognizable and nonsensical inputs, such as "i i i i i i i i [...]," it still tries to provide a fluent translation – and ends up "hallucinating" sentences.

Why religious texts? Since religious texts like the Bible and the Qur'an exist in many languages, they probably make up a large portion of the available training data for translations to or from low-resource languages.

One solution for improving translation accuracy for low-resource language pairs is to go through a third language. For example, the training data between Hungarian and Igbo may be too scarce to result in a reasonably performing translation model. However, there is plenty of accurate training data for Hungarian–English translations, and just enough for English–Igbo training. So instead of aiming for a direct Hungarian–Igbo translation, the translator would first translate from Hungarian to English and then from English to Igbo.

While this is a reasonable solution, it increases the risk for meanings getting lost in translation. As a student, I had an assignment in machine translation class in which I implemented a "bad translator." The bad translator receives an English text and translates it back to English through a chain of random languages, such as English to Czech to Swahili to Arabic to Hindu to English. Due to propagating errors, the output is sometimes nonsensical or completely different from the input. This is what I got by inputting some of the ten commandments:

> "Thou shalt not kill" was translated to "You must remove."
> "Thou shalt not make unto thee any graven image" to "You can move the portrait."
> "Thou shalt not commit adultery" to "Because you're here, try three."
> And "Thou shalt not steal" to "Woman."[1]

[1] You can test the Bad Translator at: www.cs.ubc.ca/~vshwartz/resources/bad_translator.html.

In sum, automatic translation tools like Google Translate have improved immensely in recent years. Although they are very useful, they don't work equally well for every pair of languages and every genre and topic. Blindly relying on automatic translation can cause embarrassment and misunderstanding. For this reason, automatic translation doesn't yet make second language acquisition obsolete.

Thinking in Your Native Language Makes You Sound Foreign

Mastering a second language means being able to think in that language rather than translating your thoughts from your native language. The language of our thoughts affects our word choice and grammatical constructions, so going through another language might result in incorrect or unnatural sentences.

Let me give you an example from my native language: Hebrew. I read in an online article the imperative phrase "Do sports and eat balanced." While this is understandable, it doesn't sound right in English. Given that the author had an Israeli name, I could easily reverse-engineer the English sentence and reconstruct their Hebrew thoughts.

First, the sentence was missing a noun. It should have read "eat a balanced diet." In Hebrew, omitting the noun is common practice, and the word "diet" is implied. In English, an adjective such as "balanced" can only modify a noun. If "balanced" was meant to modify the verb "eat," it should have been an adverb, but I don't think that "balancedly" is a word. Second, the word choice is odd because in English the word "sports" typically refers to competitive sports. The author likely meant "fitness" or "exercise," which also translates to "sport" in Hebrew. Finally, starting a sentence with the word "do" might prompt the reader to look for a question, as in "Do sports and eat balanced . . .?" This would be less confusing in speech, when the speaker can emphasize different words in the sentence to convey the intended grammatical role of "do." An emphasis on "do" implies an imperative, whereas an emphasis on "sports" implies a question. Either way, a literal translation of this weird sentence back to Hebrew reads perfectly normally.

While I could be smug about noticing other people's errors, my English is not unaffected by Hebrew. I once used the phrase "private case" instead of "special case" in an academic paper, literally translating the corresponding Hebrew phrase. It went unnoticed by my coauthors, one of whom was a native English speaker, but was later pointed out by another Israeli researcher.