

Error Tolerance of Machine Translation: Findings from Failed Teaching Design¹ (机器翻译的容错性: 从一个失败的教学设计中得到意外发现)

Tian, Ye
(田野)

University of Pennsylvania
(宾夕法尼亚大学)
tianyel@sas.upenn.edu

Abstract: Science is constantly being revised, and failure is crucial to this process. This study is enlightened by the findings of a failed teaching design that aims to train intermediate level Chinese language learners to use Machine Translation as a self-editing tool to improve their writing proficiency. This study finds that Sogou Translate possesses a powerful error tolerance feature that can correctly translate students' Chinese sentences into correct English sentences, even though these sentences may contain various errors made by Chinese language learners, including misspellings, typos, misusing homonyms, misusing the part of speech of a word, misusing synonyms, ungrammatical phrases, and incorrect word or sentence order. This study also suggests three areas to further investigate and to apply error tolerance features of Machine Translation in the field of Chinese language teaching.

摘要: 失败对科学的发展至关重要, 本研究同样源自一项不成功的中文教学设计。该教学设计训练中级水平的中文学生使用机器翻译作为自主纠错工具, 目的是为了他们的写作水平。然而本研究发现搜狗翻译包含强大的容错性, 即使这些学生的中文写作中包含学习者常犯的各种错误, 搜狗翻译仍然可以将其翻译成正确的英文。这些错误包括拼写错误, 写错别字, 误用同音异义词, 误用词类, 误用同义词, 不合语法的表达, 错误的词序语序等。本研究亦指出在中文教学领域进一步研究甚至应用机器翻译容错性的三个方向。

Keywords: Machine Translation, Sogou Translate, error tolerance, Chinese language teaching, self-editing

¹ This work was supported by a SAS Language Teaching Innovation Grant, Penn Language Center, University of Pennsylvania.

关键词：机器翻译、搜狗翻译、容错度、中文教学、自主纠错

1. Introduction

Computer-Assisted Language Learning (CALL) developed rapidly in the past two decades. Studies have shown that digital teaching tools, such as Machine Translation, Speech Synthesis, Speech Recognition, and Online Testing Solutions, have great potential to improve students' language proficiency levels in this mobile era. If used properly, Machine Translation can be used to enhance learners' awareness of self-correction on grammar, pronunciation, and word choice errors (Baker, 2013; Case, 2015; Correa, 2014; Groves & Mundt, 2015; Niño, 2009).

Machine Translation is not a new technology. The concept of Machine Translation emerged as early as the 1950s. The first successful large-scale application of Machine Translation was the Logos Machine Translation System, which translated English military manuals into Vietnamese during the Vietnam War in the 1970s. Unfortunately, restricted by unsatisfactory accuracy and expensive equipment, Machine Translation did not really impact foreign language teaching until Google Translate launched its phrase-based statistical machine translation (SMT) service in 2006 (Garcia & Pena, 2011; Somers, 2003).

The exact process is a patented secret but, roughly speaking, this is how Google's SMT-based algorithm works: Google collected billions of parallel documents that have been translated by human translators, such as official communications issued by the United Nations and the European Union, records of international tribunals, international company reports, and articles and books in bilingual forms. The principle is that the text one wants to translate may have been translated before by humans and can be found in their parallel text corpora. The algorithm of SMT analyzes the source language, detects the patterns, calculates the probabilities, and then generates the most likely translation in the target language (Benjamin, 2019; Grajales, 2015).

Although the accuracy of SMT has sometimes been questioned and mocked by human translators or the native speakers of the target languages, Google Translate evolved significantly along with the rapid development of information technology, network technology, and machine learning in the past decade. Google Translate soon developed into a neural network-powered machine translation engine, Neural Machine Translation (NMT), which is an end-to-end learning approach for automated translation. When it progressed into Google's Neural Machine Translation system (GNMT) in 2016, researchers claimed that:

Human evaluations show that GNMT has reduced translation errors by 60% compared to our previous phrase-based system on many pairs of languages: English ↔ French, English ↔ Spanish, and English ↔ Chinese. Additional experiments suggest the quality of the resulting translation system gets closer to that of average human translators (Wu et al., 2016, p. 1).

The detailed algorithms of GNMT are beyond the scope of this study. Explaining their jargon, such as “LSTM network,” “residual connections,” “attention mechanism,” “the decoder/encoder,” “low-precision arithmetic,” “inference computations,” “beam search technique,” “length-normalization procedure,” or “coverage penalty” will not break open the black box of GNMT. But a statement made by Barak Turovsky (2016), the Product Lead of Google Translate, might shed some light on the mystery.

At a high level, the Neural system translates whole sentences at a time, rather than just piece by piece. It uses this broader context to help it figure out the most relevant translation, which it then rearranges and adjusts to be more like a human speaking with proper grammar. Since it's easier to understand each sentence, translated paragraphs and articles are a lot smoother and easier to read. And this is all possible because of end-to-end learning system built on Neural Machine Translation, which basically means that the system learns over time to create better, more natural translations (para. 2).

Besides Google Translate, many technology companies have developed and launched their own Machine Translation services and products, such as Microsoft Translate Service, Siri Translate, Translate Facebook, DeepL Translator, Baidu Translate, Sogou Translate, and WeChat Translate. All of them not only offer website interfaces, but also mobile apps for Android and iOS. Thus, with the widespread use of smartphones, Machine Translation has become a double-edged sword in the field of foreign language education: while cheating is easier with the advent of the technology, it also has potential to become a valuable pedagogical tool.

While many foreign language instructors, as indicated in the following section, have claimed they successfully included Machine Translation into teaching practice, this study introduces a failed teaching design that aims to incorporate and evaluate Machine Translation as a means of Computer-Assisted Language Learning (CALL) for teaching in the Mandarin classroom. As with every major scientific innovation, there are countless failed projects that are equally important to the advancement of science. Although this teaching design did not achieve its original goals, it accidentally generated other insightful findings about the error tolerance of Machine Translation, which may enlighten future research and teaching practice.

2. Literature Review

Most of the previous research on Machine Translation in the field of foreign language teaching not surprisingly focuses on Google Translate, which is the most available Machine Translation tool for most languages. Related research mainly centers on the following three categories. The first investigates, of course, the accuracy of Google Translate. For example, Correa (2014) introduces some of the advantages and disadvantages of English \rightleftharpoons Spanish translation in Google Translate. She states that Google Translate is “good at conjugating, spelling, basic agreement, and some common idioms”

(Correa, 2014, p. 7). But Google Translate often results in literal translation or unnatural writing that contains grammatical inaccuracies, discursive inaccuracies, and errors that humans do not commit. Google Translate is often “unable to account for cultural references and other extra-linguistic issues such as context, connotation, denotation or register” (Correa 2014, p. 7) and has difficulty with some idioms. Google Translate often does not translate misspelled words or proper nouns in the original text (they are just reproduced). Groves and Mundt (2015) point out that the accuracy of Google translations is close to the intermediate level of English learners when translating Malay and Chinese to English, and Google’s accuracy improves over time.

The second category usually surveys the way in which foreign language learners use Machine Translation and the various (but more often diametrically opposed) opinions of foreign language instructors and learners toward the use of Machine Translation. Many studies have shown that students use Machine Translation even though their instructors explicitly prohibit it (Clifford, Merschel, & Munné, 2013; Correa, 2011, 2014; García & Pena, 2011; Niño, 2009). For example, Clifford, Merschel, and Munné (2013) use a questionnaire to survey instructors and students in the Department of Romance Studies at Duke University. They find that instructors and students have different attitudes toward Machine Translation. Although students are aware that Machine Translation is imperfect, more than 88% of respondents admit that they have used Machine Translation in their studies, and most students believe that Machine Translation helps them learn new languages. But most of the instructors they surveyed are skeptical about the effectiveness of Machine Translation. They believe that the emergence of Machine Translation interferes with traditional classroom teaching. Therefore, students are prohibited from using Machine Translation to finish assignments in the syllabus. It is worth noting that not all foreign language instructors object to the use of Machine Translation. For example, Niño (2009) and Baker (2013) show that there are still some foreign language instructors and students who maintain a positive attitude toward Machine Translation, especially instructors who are devoted to computer-assisted teaching and students in upper-level courses. They believe that the shortcomings of Machine Translation can even be used to improve students’ understanding of the complexity of translation and language learning.

The third category discusses how to incorporate Machine Translation in foreign language classrooms to improve students’ language proficiency. For instance, Correa (2014) states that since the output of Google Translate is far from perfect, Spanish instructors can have their students decipher the intended meaning and edit the output of Google Translate as a post-editing exercise. And as a pre-editing exercise, “students can pre-edit a text that they wrote in Spanish until they get an acceptable translation in English” (Correa, 2014, p. 11), which “can help with, among other issues, accent placement or gender assignment” (Correa, 2014, p. 12). In fact, Case (2015) summarizes that many studies demonstrate how to utilize the imperfect translation that is generated by Machine Translation for students to conduct pre-editing or error correction. Similar teaching designs can be found in Groves & Mundt (2015), Kliffer (2005), Niño (2008), Somers (2003), and in Zanettin (2009). This kind of research is more meaningful to foreign language teaching since it is directly related to classroom instruction.

Compared with other languages, studies on Machine Translation in Chinese language teaching and learning are rather inadequate in North America. Only Tian (2018) has published a study, which analyzes about 500 translation exercises that are collected from students' homework. He finds that students rely on Google Translate to do their homework in varying degrees. The accuracy of Google Translate is not unacceptable because Google Translate can do a better job than intermediate-level Chinese language learners, and it evolves very quickly. In terms of Chinese ⇌ English translation, Sogou Translate [搜狗翻译] and Baidu Translate [百度翻译] are more accurate than Google Translate. This research shows that Machine Translation is accurate enough for students to use it as a shortcut to achieve a good grade.

3. An Interesting Failed Teaching Design

The failed teaching design introduced in this study is inspired by previous research that uses Machine Translate as a self-editing tool in foreign language classrooms. It is conducted in an intermediate (third-year) level Chinese language course at an Ivy League university, which uses *A New China* (《新的中国》) as its textbook. This design aims to help students gain experience in becoming self-assessors via Sogou Translate when they do their homework. Students are required to write an essay in Chinese first, then use Sogou Translate to translate their Chinese writings into English to check if their Chinese writings contain any obvious mistakes.² The intended objective is that since Sogou Translate is very accurate at least at the intermediate level, if the English translation looks incorrect to students, then students' original Chinese sentences are probably wrong. Students can go back to revise their Chinese essay until they get an acceptable English translation. Example 1 is instruction from one of the assignments. The writing sections of all twenty assignments for the entire academic year have been modified based on this design.

Example 1

Please write a short paragraph to introduce this image in Chinese, and then record your answer. Use as many new words or patterns as possible.

place+有+.....; place+到处都 verb; 各种各样+的; 教+sb.+ verb /怎么 verb;, 有的....., 有的....., 最常见的是.....; 和/跟.....有关的; 和/跟 A 比(起来), B 更/比较 adj; 大大地+verb; 带着+noun; subj. (向 sb.) + 提出(了)+noun; 改掉; 乱+verb; 随便+verb; 值得;



Requirements and Tips:

² This paper does not make a distinction between “error” and “mistake” from a linguistic viewpoint. It uses the them interchangeably.

- A. Write down your answer in Chinese. (250 to 300 characters. **Tips: Please utilize Sogou Translate to check if your answer is correct.**)
- B. Record your answer. (Tips: You can utilize the embedded “read” function of Sogou Translate to practice your pronunciation).
- C. Upload both your writing and your recording here.

However, this study finds that the above teaching design does not really achieve its intended goal. Students’ Chinese writings still contain many obvious errors, such as ungrammatical expressions, incorrect word order, misspellings, malapropisms, homonyms, and typos, even though students used Sogou Translate to check their answers before they submitted their homework. It turns out that the error tolerance feature embedded in Machine Translation is so advanced that it can translate those mistakes made in students’ Chinese writing into acceptable English. Thus, students would assume that their original Chinese writings are also correct. In fact, when designing similar pre-editing exercises via Google Translate in her Spanish course, Correa (2014) warns that as Machine Translate tools

are getting better at guessing missing accents and reassigning gender, the possibility exists that an ungrammatical input ... could produce a grammatical output. For this reason, it may be wise to remind students that having a good translation is never a guarantee that the original text is also error-free (or vice versa) (p. 12).

However, this warning is overlooked by many previous, similar studies, including the aforementioned teaching design, because this phenomenon has never been systematically examined before.

4. Error Tolerance of Sogou Translation

Although the teaching design does not achieve its intended goal, it generates equally or maybe more interesting findings. This study discovers that when translating Chinese to English, Sogou Translate is very tolerant of errors in the original Chinese text. In other words, even though students’ Chinese writings may contain assorted mistakes that Chinese language learners can make, Sogou Translate can still “guess” the “correct” meanings and translate them into acceptable English.

Error tolerance is the design of things via various technologies “to be resilient to human error.” “When a human error does occur, error-tolerance design gracefully detects and handles it” (Spacy, 2017, p. 1). One of the most common error tolerance features might be input correction. For example, Microsoft Word’s AutoCorrect feature can automatically fix misspelled words or correct capitalization of words. And Siri’s voice commands may accept different variations of the same command. In the case of Sogou Translate, the error tolerance feature allows the machine to accept imperfect Chinese inputs and translate them into English. This is a basic and necessary design because in real-life scenarios even native Chinese speakers cannot guarantee that they will speak flawless or grammatically correct Chinese sentences all the time. This research reveals that the error tolerance feature of

Sogou Translate can ignore at least the following four types of common mistakes made by Chinese language learners, which usually would not be acceptable by native Chinese speakers or Chinese language instructors.

4.1 Misspelling Errors

When Chinese language learners type Chinese characters, one common error is that they misspell the Pinyin and then choose the wrong characters (or correctly type the pinyin but still choose the wrong characters). However, this study shows that Sogou Translate can detect the mistyped words due to similar sounds and guess the intended meaning based on the context and then generate the correct English translation. For instance, although a student misspelled 将来 (*jianglai*, which means “in the future”) as 经来 (*jinglai*, which is not a Chinese word at all) in Example 4.1.1, Sogou Translate still translated it into “in the future” correctly. And in Example 4.1.2, when a student misspelled 奇怪 (*qiguai*, which means “strange”) into 吸怪 (*xiguai*, which is not a Chinese word at all), Sogou Translate also correctly translated it. Similar phenomenon can also be found in Example 4.1.3 where 开个开放 was correctly (and surprisingly) translated into “opening up.”

Example 4.1.1

Student’s incorrect sentence:

我觉得经来，美国人应该欢迎移动支付，要不然不能跟中国竞争。

Sogou Translate result (August 1st, 2019):

I think Americans should welcome mobile payment **in the future**, otherwise they cannot compete with China.

Example 4.1.2

Student’s incorrect sentence:

视频里的一个人告诉了主持用现金可能会给北京人一个吸怪的印象。

Sogou Translate result (August 1st, 2019):

One of the people in the video told the host that using cash might give Beijingers a **strange** impression.

Example 4.1.3

Student’s incorrect sentence:

开个开放以后中国不再是铁饭碗的城市。

Sogou Translate result (January 8, 2020):

After **opening up**, China is no longer an iron rice bowl city.

4.2 Vocabulary Errors

Messing up a word’s part of speech or using the wrong synonyms are also two common mistakes that Chinese language learners may struggle with when learning Chinese vocabulary. When dealing with Chinese sentences with such errors, Sogou Translate does not conduct a simple word-for-word translation and then produce an English sentence with

errors; instead, it fixes the errors “behind the scenes” first and then generates the correct translation. For instance, in Example 4.2.1, a student incorrectly wrote 浪费花钱 (because 浪费 should be followed by a noun, not a verb), but Sogou Translate still correctly translated into “is wasting its money” not “is wasting spending money.” In example 4.2.2, a student confused the meanings of 肯定 (which expresses the certainty of speculation) and 当然 (which expresses the certainty of a fact) in his Chinese sentence, but Sogou Translate still correctly translated it. In example 4.2.3, Sogou Translate first corrects a student’s misspelling error (it should be 名字 not 明子) by ignoring that mistake, and then correctly translates the misused noun (意思 which means “meaning”) into the verb “means.” And in example 4.2.4, a student mixed up the meanings of 认为 (“to believe”) and 想起 (“to think of, to remind”) in her Chinese sentence, but Sogou Translate still correctly translated it.

Example 4.2.1

Student’s incorrect sentence:

现在，美国的交通和经济很不好。他说这是因为美国浪费花钱在中东，所以现在没有现代化的交通工具。

Sogou Translate result (August 8, 2019):

At present, the transportation and economy in the United States are very bad. He said this is because the United States is **wasting its money** in the Middle East, so there is no modern means of transportation.

Example 4.2.2

Student’s incorrect sentence:

我去过中国也看到了他们的铁路，机场，和别的基础设施建设，肯定比美国漂亮和有效了。

Sogou Translate result (January 8, 2020):

I have been to China and have seen their railways, airports and other infrastructure construction, which are **certainly** more beautiful and effective than those in the United States.

Example 4.2.3

Student’s incorrect sentence:

“Rust Belt” 的意思是很老的工厂地方。一方面这个明子意思工厂地方，一方面这意思钢城市。

Sogou Translate result (January 14, 2020):

Rust Belt means a very old factory place. On the one hand, it **means** a factory, on the other hand, it **means** a steel city.

Example 4.2.4

Student’s incorrect sentence:

以前在美国这个名字很有名。现在这个名字让我认为幽灵城市。

Sogou Translate result (January 14, 2020):

The name used to be very famous in America. Now the name makes me **think of** ghost city.

4.3 Grammar Errors

This study shows that Sogou Translate can also “overlook” some of the grammatical errors that exist in students’ Chinese writings, guess the intended meaning, and then produce the correct English translation, which is a capability that many novice Chinese language instructors might not even possess. For instance, in Example 4.3.1, although a student wrote an incorrect comparative bi-structure (比字句) that is, frankly speaking, difficult to understand even for an experienced Chinese language instructor, Sogou Translate produced the correct English sentence nonetheless. Similar grammar corrections can also be found in Example 4.3.2, Example 4.3.3, and Example 4.3.4, in which the student mistakenly used lian-structure (连字句), rang-structure (让字句), put a comma before the conjunction word 和 [and], and connected two adjectives with the conjunction word 和 respectively.

Example 4.3.1

Student’s incorrect sentence:

但是在美国，人比手机多用电脑。

Sogou Translate result (July 14, 2019):

But in America, people use computers more than mobile phones.

Example 4.3.2

Student’s incorrect sentence:

比如说新加坡的机场连有电影院。

Sogou Translate result (January 8, 2020):

For example, Singapore’s airport has cinemas.

Example 4.3.3

Student’s incorrect sentence:

可是在政府之内总是有冲突，让政府不取得很大进步

Sogou Translate result (January 8, 2020):

However, there are always conflicts within the government, which prevents the government from making great progress.

Example 4.3.4

Student’s incorrect sentence:

我去过中国也看到了他们的铁路，机场，和别的基础设施建设，肯定比美国漂亮和有效了

Sogou Translate result (January 8, 2020):

I have been to China and have seen their railways, airports and other infrastructure construction, which are certainly more beautiful and effective than those in the United States.

4.4 Word and Sentence Order Errors

Another typical mistake often made by Chinese language learners who are native English speakers is getting word and sentence order wrong, which is caused by the negative transfer between Chinese and English. For instance, they tend to put time and location at the end of the sentence, which is incorrect in Chinese. In Example 4.4.1, a student mistakenly put the location words 在中东 [in the Middle East] at the end of a Chinese sentence, but Sogou Translate rendered it into correct English. Such is the case with Example 4.4.2 about the time phrase.

Example 4.4.1

Student's incorrect sentence:

现在，美国的交通和经济很不好。他说这是因为美国浪费花钱在中东，所以现在没有现代化的交通工具。

Sogou Translate result (August 8, 2019):

At present, the transportation and economy in the United States are very bad. He said this is because the United States is wasting its money **in the Middle East**, so there is no modern means of transportation.

Example 4.4.2

Student's incorrect sentence:

美国政府把许多的工厂设在东北部在 60 年代。

Sogou Translate result (January 8, 2020):

The U.S. government set up many factories in the northeast **in the 1960s**.

In addition, when composing a Predicate Adjectives sentence (形容词谓语句), which means using an adjective as a verb/predicate in a sentence (e.g. 学中文很麻烦), influenced by their mother tongue, English speakers sometimes mistakenly put the adjectives at the beginning of the Chinese sentence. But once again, Sogou Translate can produce a beautiful and grammatical English sentence from a student's ill-formed Chinese sentence. Example 4.4.3 and Example 4.4.4 illustrate this phenomenon.

Example 4.4.3

Student's incorrect sentence:

这是因为很贵建设新的基础设施建设

Sogou Translate result (July 14, 2019):

This is because it is very expensive to build new infrastructure

Example 4.4.4

Student's incorrect sentence:

很麻烦得到允许在私人的土地建设。

Sogou Translate result (July 14, 2019):

It is very troublesome to get permission to build on private land.

Sogou Translate can also fix similar word order mistakes when students put professional titles before a person's name in Chinese, such as the Example 4.4.5.

Example 4.4.5

Student's incorrect sentence:

总统特朗普

Sogou Translate result (July 14, 2019):

President Trump

5. Discussion

Besides the error tolerance of Sogou Translate, this study also discovers two more interesting findings that are worth further discussion.

5.1 Error Tolerance of Google Translate

Tian (2018) shows that Sogou Translate is more accurate than Google Translate in terms of English \rightleftharpoons Chinese translation. This research provides new evidence for this conclusion because the function of error tolerance of Sogou Translate is more powerful than Google Translate. In other words, when the original Chinese input is imperfect, the same as Sogou Translate, Google Translate can also sometimes guess the intended meaning and generate the correct English translations, but Sogou Translate does a better job than Google Translate. For instance, Google Translate fails to decipher the real meaning of 经来 (*jinglai*, which is not a Chinese word) in Example 5.1.1. It chooses not to translate it at all. Similar evidence that Sogou Translate is better than Google Translate can also be found in Example 5.1.2, Example 5.1.3, and Example 5.1.4.

Example 5.1.1

Student's incorrect sentence:

我觉得经来，美国人应该欢迎移动支付，要不然不能跟中国竞争。

Sogou Translate result (August 1, 2019):

I think Americans should welcome mobile payment **in the future**, otherwise they cannot compete with China.

Google Translate result, (January 11, 2020):

In my opinion, Americans should welcome mobile payments, or they cannot compete with China.

Example 5.1.2:

Student's incorrect sentence:

开个开放以后中国不再是铁饭碗的城市。

Sogou Translate result (January 8, 2020):

After **opening up**, China is no longer an iron rice bowl city.

Google Translate result, (January 11, 2020):

After **opening**, China is no longer a city of iron rice bowls.

Example 5.1.3:**Student's incorrect sentence:**

可是在政府之内总是有冲突，让政府不取得很大进步

Sogou Translate result (January 8, 2020):

However, there are always conflicts within the government, **which prevents** the government from making great progress.

Google Translate result, (January 11, 2020):

However, there are always conflicts within the government, **so that** the government does not make much progress

Example 5.1.4:**Student's incorrect sentence:**

很麻烦得到允许在私人的土地建设。

Sogou Translate result (July 14, 2019):

It is very troublesome to get permission to build on private land.

Google Translate result, (January 11, 2020):

Very trouble getting permission to build on private land.

5.2 “Two-Step Translation” Method

Trying to guess students' intended meaning and/or correcting students' ill-formed Chinese sentences (e.g. Example 4.3.1 “但是在美国，人比手机多用电脑”) can sometimes be very challenging even for experienced Chinese language instructors. But the error tolerance feature of Sogou Translate may help instructors by ignoring those mistakes and getting the correct English translation. Instructors can simply use Sogou Translate to translate that English sentence back into Chinese and possibly get a correct (and often a better) Chinese sentence that the student wants to express. Examples 5.2.1, 5.2.2, and 5.2.3 show how a student's incorrect and hard-to-understand Chinese sentence is translated into a flawless Chinese sentence via this “two-step translation” method.

Example 5.2.1**Student's incorrect sentence:**

但是在美国，人比手机多用电脑。

Sogou Translate result (from Chinese sentence to English) (January 8, 2020):

But in America, people use computers more than mobile phones.

Sogou Translate result (from English sentence to Chinese) (January 8, 2020):

但是在美国，人们使用电脑多于手机。

Example 5.2.2**Student's incorrect sentence:**

第二个事的引述是川普在发言说的。

Sogou Translate result (from Chinese sentence to English) (January 8, 2020):

The second thing is quoted by Trump in his speech.

Sogou Translate result (from English sentence to Chinese) (January 8, 2020):

川普在演讲中引用了第二点。

Example 5.2.3

Student's incorrect sentence:

曾经是我完全成为天主教徒的时候，我这样做了。

Sogou Translate result (from Chinese sentence to English) (January 30, 2020):

It was when I became a full Catholic that I did so.

Sogou Translate result (from English sentence to Chinese) (January 30, 2020):

我是在成为一名正式天主教徒后才这样做的。

Moreover, this study also finds that the error tolerance feature of Sogou Translate probably works better on longer passages than shorter sentences, which means the more context provided, the more accurate Sogou Translate is. For instance, Example 5.2.4 is a student's essay explaining the term "Rust Belt" in Chinese, which is poorly composed and consists of many errors. When using Sogou Translate to translate a single ill-formed sentence in that essay, "一方面这个明子意思工厂地方," it gets an equally strange English sentence "on the one hand, this place where Akiko means factory." However, when translating the whole paragraph, Sogou Translate renders a highly acceptable English paragraph, and many errors, including that sentence, in the original text, are correctly translated. And Sogou Translate can generate a much better Chinese paragraph (only containing a few minor errors) by translating this English paragraph back into Chinese again.

Example 5.2.4

Student's incorrect sentence:

一方面这个明子意思 工厂地方。

Sogou Translate result of this single sentence (January 15, 2020):

On the one hand, this place where Akiko means factory.

Student's poorly composed essay:

"Rust Belt" 的意思是很老的工厂地方。一方面这个明子意思工厂地方,一方面这意思钢城市。这个现象发生了从十九世纪中到二十世纪中 在美国的北部。美国用很大的地方来造成一个产业地方。当时美国政府送很多技术给这个城市来造成一个重要的产业城市。以前在美国这个名字很有名。现在这个名字让我认为幽灵城市。以前这地方是一个成功的产业所在和当时产业上最发展的所在这主要是因为这个所在 既有丰富的资源又有廉价的劳动力。美国政府吧许多的工厂设在东北部的美国。现在这个地方减少就业机会。很多人失业了。"Rust Belt",一方面对经济做出了贡献,但是另一方面创造工人失业。那里的人们现在不享受美好的过去。

Sogou Translate result of the whole paragraph (January 15, 2020):

Rust Belt means a very old factory place. **On the one hand, it means a factory**, on the other hand, it means a steel city. This phenomenon occurred in the north of the United States from the middle of the 19th century to the middle of the 20th century. The United States uses a large area to create an industrial area. At that time, the U.S. government sent many technologies to the city to create an important industrial city. The name used to be very famous in America. Now the name makes me think of ghost city. In the past, this place was the place where a successful industry was located and where the industry was most developed at that time, mainly because this place has rich resources and cheap labor. The U.S. government has many factories in the northeastern United States. This place is now reducing employment opportunities. Many people have lost their jobs. Rust Belt, on the one hand, has contributed to the economy, but on the other hand, has created unemployment among workers. The people there do not enjoy the beautiful past now.

Sogou Translate result (from English sentence to Chinese) (January 15, 2020):

锈带意味着一个非常古老的工厂。一方面，它意味着一个工厂，另一方面，它意味着一个钢铁城市。这种现象发生在19世纪中叶至20世纪中叶的美国北部。美国使用大面积来创建工业区。当时，美国政府向这座城市发送了许多技术，以创建一个重要的工业城市。这个名字过去在美国非常有名。现在这个名字让我想起了鬼城。过去，这个地方是一个成功产业的所在地，也是当时该产业最发达的地方，主要是因为这个地方资源丰富，劳动力便宜。美国政府在美国东北部有许多工厂。这个地方现在正在减少就业机会。许多人失业了。锈带一方面促进了经济发展，但另一方面也造成了工人失业。那里的人们现在不喜欢美好的过去。

6. Pedagogical Implications and Recommendations for Future Research

Science cannot succeed without failures (Firestein, 2015). The findings from the aforementioned unsuccessful teaching design do not indicate that Machine Translate cannot be applied in the Chinese language classroom but only mean that using Machine Translate to design pre-editing exercises needs more sophisticated consideration. Moreover, this study suggests that there are at least three possibilities to further investigate how to utilize Machine Translate in Chinese language teaching.

1) This study is a qualitative case study of only one intermediate-level Chinese language course, which only investigates the error tolerance in Sogou Translate when conducting Chinese- to-English translation. Student samples and their Chinese writing examples do not reflect any quantitative statistical significance. Future researchers can further explore the phenomena of error tolerance of Machine Translation by expanding their research samples to include students at different language proficiency levels and

investigating similar error tolerance phenomena that may exist in English-to-Chinese translation.

2) Most of the Machine Translation tools, such as Google Translate, Sogou Translate, and Baidu Translate, offer speech recognition features, which are also tolerant of pronunciation errors (e.g. different accents). This means that the speech recognition function can correctly recognize speech with some flaws (Ruiz, Bertoldi, & Federico, 2019). In teaching Chinese language, instructors often highlight the importance of the accuracy of students' pronunciation, which is mainly judged by ear. How accurate is accurate enough? There is a lack of direct and objective standards. Is "machine-acceptable" accuracy good enough? In the future, a study of error tolerance that uses Chinese language learners' speech as research samples will be very promising.

3) Future research may focus on investigating how to train Chinese language learners to become self-assessors and make progress on their writing proficiency without the instructor's involvement via the aforementioned "two-step translation" method. Computer engineers may even develop an "automatic composition revision" program via this editing technique.

7. Conclusion

Science constantly benefits from failures. Failures often provoke very unpredictable insights and compel researchers to look at a problem differently. This study is inspired by an unsuccessful teaching design that uses Sogou Translate as a self-editing technique. It finds that when conducting Chinese-to-English translation, Sogou Translate possesses a powerful error tolerance function. It can correctly translate erroneous Chinese sentences composed by intermediate level Chinese language learners into grammatical English sentences. These errors include misspelling, misusing homonyms, typos, misusing the part of speech of a word, misusing synonyms, ungrammatical phrases, and incorrect word/sentence order. The error tolerance function in Sogou Translate is more powerful than Google Translate, which proves again that Sogou Translate is a better tool than Google Translate in terms of Chinese \rightleftharpoons English translation. This study also suggests three possible areas for future research on Machine Translation. These include: investigating error tolerant algorithms that exist in English to Chinese translation; examining to what extent Machine Translation can tolerate a Chinese language learner's imperfect pronunciation, and; exploring how to train students to gain experience in self-assessing their oral and written Chinese via Machine Translation. These opportunities might raise students' metalinguistic awareness and help prepare them for lifelong learning.

References

- Baker, C. L. (2013). *Student and instructor perceptions of the use of online translators in English composition* (Order No. 1536057). Available from ProQuest Dissertations & Theses Global. (1352759991).

- Benjamin, M. (2019). Teach you backwards: An in-depth study of Google Translate for 103 languages. Retrieved from <https://teachyoubackwards.com>
- Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *eLearning Papers*, (45), 4-16.
- Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *The Acquisition of Second Languages and Innovative Pedagogies*, (10), 14. doi:10.7203/attic.10.2228
- Correa, M. (2011). Academic dishonesty in the second language classroom: Instructors' perspective. *Modern Journal of Language Teaching Methods*, 1(1), 14.
- Correa, M. (2014). Leaving the "peer" out of peer-editing: Online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin American Journal of Content & Language Integrated Learning*, 7(1), 1-20. doi:10.5294/laclil.2014.7.1.1
- Firestein, S. (2015). *Failure: Why science is so successful*. Oxford University Press.
- Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning*, 24(5), 471-487. doi:10.1080/09588221.2011.582687
- Grajales, C. A. G. (2015). The statistics behind Google Translate. Retrieved from <https://www.statisticsviews.com/details/feature/8065581/The-statistics-behind-Google-Translate.html>
- Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, 37, 112-121. doi:10.1016/j.esp.2014.09.001
- Kliffer, M. (2005). *An experiment in MT post-editing by a class of intermediate/advanced French majors*. 160-165. Paper presented at the 10th Annual Conference of the European Association for Machine Translation, Budapest, Hungary.
- Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, 21(1), 29-49. doi:10.1080/09588220701865482
- Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241-258. doi:10.1017/S0958344009000172
- Ruiz, N., Bertoldi, N., & Federico, M. (2019). Assessing the tolerance of neural machine translation systems against speech recognition errors. Retrieved from <https://arxiv.org/pdf/1904.10997v1.pdf>. doi:10.21437/Interspeech.2017-1690
- Somers, H. L. (2003). *Computers and translation: A translator's guide*. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Spacey, J. (2017). 10 Examples of error tolerance. Retrieved from <https://simplicable.com/new/error-tolerance>
- Turovsky, B. (2016). Found in translation: More accurate, fluent sentences in Google Translate. Retrieved from <https://www.blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/>
- Tian, Y. (2018). The challenge of machine translation to traditional translation homework in Chinese language learning. *Journal of Technology and Chinese Language Teaching*, 9(1), 78-95. Retrieved from

- <http://www.tclt.us/journal/2018v9n1/tian.pdf>. [田野. 2018. 机器翻译对中文传统翻译作业的挑战. 科技与中文教学, 9(1), 78-95.]
- Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., . . . Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. <https://arxiv.org/abs/1609.08144>
- Zanettin, F. (2009). Corpus-based translation activities for language learners. *The Interpreter and Translator Trainer*, 3(2), 209-224.
doi:10.1080/1750399X.2009.10798789