

On Improper Machine Translations in Press Reports*

Huiqiong Duan

School of Foreign Languages, Nanchang Hangkong University, Nanchang 330063, China

Xinyu Hu

School of Foreign Languages, Nanchang Hangkong University, Nanchang 330063, China

Yidan Gao

School of Communication, Hong Kong Baptist University, Hongkong, China

Abstract—Machine translation, also known as automatic translation, is the process of converting one natural language (source language) into another natural language (target language) by using networks. There are some language errors in current machine translation in news releases. Having compared human translators' translation texts and machine translation results, improper machine translation results are found. They are inaccurate use of words, rigid sentence patterns and unclear expression of specific cultural meanings. Accurate machine translation needs the assistance of human translators.

Index Terms—artificial intelligence, machine translation, language errors

I. INTRODUCTION

Background

The research history of machine translation can be traced back to the 1930s. In the early 1930s, the French scientist G.B. Alchouni came up with the idea of machine translation. In 1933, the Soviet inventor Trotansky designed a machine to translate one language into another. In 1949, W. Weaver published the Memorandum of Translation, which formally proposed the idea of machine translation (Tian Jiangbo, 2018). In 1954, (Georgetown University) of Georgetown University in the United States, in collaboration with IBM Company, completed the experiment of English-Russian machine translation with IBM-701 computer for the first time, demonstrating the feasibility of machine translation to the public and the scientific community. Thus, the prelude to the study of machine translation has been opened.

It is not too late for China to start the research. As early as 1956, the state included the research of machine translation in the national scientific work development plan. In 1957, the Institute of languages of the Chinese Academy of Sciences and the Institute of Computational Technology conducted experiments on Russian-Chinese machine translation, which translated nine different types of more complex sentences. From the 1950s to the first half of the 1960s, machine translation studies were on the rise. For military, political and economic purposes, the United States and the former Soviet Union had provided substantial financial support for machine translation projects. However, due to the geopolitical and economic needs, European countries have attached great importance to the research of machine translation. Machine translation has been on the upsurge for a while. In this period, machine translation has entered an optimistic period of prosperity.

Google Translation is one of machine translation websites that has many users. It is a multilingual translation of text, voice and image (Liu Qiyuan, 2016). The essence of Google translation is based on multi-language parallel corpus, combining statistical and mathematical methods, to construct big data analytical model to mine the inherent laws between different languages. Google translation is not limited by the number of words in the original text, and has a super-retrieval function, which is engaged in almost all industries of translation. In 2016, Google replaced all-product line translation algorithms with neural network-based machine translation systems, using state-of-the-art training techniques to improve translation quality (Una Zhang, 2018).

Chinese researchers focus on vocabulary and grammar research of machine translation, researching on sentence pattern errors, and political language translation, but few achievements have been made in news language translation of machine translation. For example, Lin Xiaoping (2018) published the small clause complex theory under the framework of Chinese-English machine translation error analysis (Wu, Yonghui, Schuster, Mike, Chen, Zhifeng, etc., 2016). It analyzes the characteristics of small compound sentences and translation methods, as well as the various aspects of language in machine translation errors. Studying machine translation from the perspective of computer application, Xu Xuehui (2018) from Beijing Foreign Studies University published *Chinese-English translation Quality Evaluation of*

* **Fund projects:** Jiangxi Colleges Liberal Arts Programs "Press release machine translation between English and Chinese language error analysis", "Mobile learning in college English vocabulary learners: an empirical study of metacognitive strategy training" (YY19223); Jiangxi Educational Reform Subject (JXJG - 13-8-19); Jiangxi 13.5 Planning Project (Numbers: 17ZD030) and 13.5 (17 TB118);

Neural Network Machine Translation. Jiang Baoxuan from Beijing Foreign Studies University published *Diagnostic Evaluation of Chinese-English interpretation quality of Neural Network Machine Translation* in 2016.

At present, under the influence of economic globalization, news of a country would spread widely around the world. With the help of machine translation, press releases of a country are translated into many languages. But translation errors may also exist in these influential press releases. According to *Analysis of Lexical and Syntactic Errors in English-Chinese Machine Translation* (Yin Jiao, 2017), Google, Youdao and Bing translations have 22 linguistic errors.

In order to further reduce the translation problems, it is necessary to further study the language errors in news articles. From the perspective of linguistics, this paper will study improper linguistic machine translations in press releases to find a better way to use machine translation.

II. CHARACTERISTICS OF NEWS LANGUAGE

News translation must be analyzed in combination with the specific features of news text. With the development of translation technology, news language has new features.

A. General Characteristics

The first feature is using new words and acronyms (Li Qiang, 2019). Acronyms are one of the main word-formation methods in modern English newspapers and periodicals, especially active in news headlines. Acronyms not only have the characteristics of concise word-making and convenient use, but also have strong vitality. English acronyms are formed by omitting or simplifying the original words by selecting the appropriate syllables or letters or numbers of the words. In the collection of news, English-speaking countries will involve many proper nouns, such as longer names and institution names.

The new words in news English reports are also accompanied by the rapid development of society and the emergence of new things, such as swine flu, avian flu, bird flu, SARS, AIQaida, shared unisex between men and women (Guo Jin, 2018). Acronyms are frequently used in news English. Words are refined. The amount of information is large, such as OPEC-The Organization of Petroleum Exporting Country Organization of Petroleum Exporting countries, IMF-international Monetary Fund International Monetary Fund.

The second characteristic is simple and plain (Rob Vandenberg, 2018). News is reported in easy-to-understand words. Therefore, the news sentences will usually not adopt complex sentence structures. For example, the "Located in southeastern Tibet at an average altitude of 3, 100 meters, Linzhi is now the fourth prefecture-level city in the Tibet autonomous region. (Guo Jin, 2018).".

The third characteristic is that news translation needs to be completed in a short time. What the news conveys is the real event. And the event information often has the timeliness. The core information in news reports will not exceed a few hundred words, and these hundreds of words must condense a lot of connotation and information.

B. New Characteristics in Intelligence Artificial Age

With the development of new media, the original artificial news production mode no longer adapt to the development of the times. Artificial intelligence news is based on computer programs and big data analysis, presenting the characteristics of standardized content, strong timeliness and hidden position, completely overturning the traditional news interview, writing, editing, proofreading and other links, making news production simple, personalized and time-free (Zhang Liang, 2019). News language has the following two features.

1. Novel, entertaining, and diversified expressions

On May 23, 2017 China Youth Daily on WeChat news "首战告负! 柯洁不敌阿尔法狗, 胜天半子还可能吗". In the article, using the flexible turn of the actual combat, AlphaGo, in the upper right corner, started the battle again, "AlphaGo, stirring a pool of spring water. Now, the last three disks of the game are only two, and the HAPPY ENDING..." is more popular in the article. In the end, it's also a joke that if you need to enhance a little, there may be a foreign aid to consider. The hall is still three seconds to the battlefield... "It's almost impossible to see in traditional media.

2. Endless new words

Neologisms are a microcosm of the changes of the times and social progress. New words emerge in endlessly, creating new examples in English news reports. After being imitated and reaching a considerable degree of universality, they are gradually absorbed into the dictionary and officially become part of modern English. For example:

Kidult (a blend of kid and adult) refers to a so-called grown-up who doesn't want to grow up (or at least act like an adult) and would instead prefer so-called "children's" stuff for entertainment, like cartoons, toys, comic books, Disney movie s, etc. (The Economist, 2014, Political Friends, Religiously Different)

A mixture of Kidult (kid and adult) refers to people who appear to be adults but don't want to grow up (or at least don't want to act like adults), who are more likely to choose children's entertainment, Such as cartoons, toys, comic books and Disney movies.

III. LANGUAGE ERRORS IN MACHINE TRANSLATION

A. Low Accuracy in Vocabulary Translation

Many words express the same meaning, while the same word, depending on the word order and the speaker's tone of voice, can represent the opposite meaning. It is an unqualified challenge for intelligent translation machines programmed by logical judgment. Let's take an article from Agence France-Presse (Agence France-Presse) on May 16, 2019 as an example.

The team, led by FIU biologist Matthew DeGennaro, identified the guilty receptor as Ionotropic Receptor 8a, or simply IR8a.

FIU is the abbreviation of Florida International University. And IRU8a originates from the abbreviation of a scientific and technological term. Therefore, when Chinese-English translation, we should deal with the understanding error caused by acronyms. In the process of translating acronyms into Chinese, the full name of acronyms should be translated and sometimes further explained. The following table is a comparison between Jinshan translation and manual translation of the previous paragraph (MAO Xiaoli, 2019).

原文	人工翻译	金山翻译
The team, led by FIU biologist Matthew DeGennaro, identified the guilty receptor as Ionotropic Receptor 8a, or simply IR8a, through a process of elimination that began in 2013 when DeGennaro was able to create the world's first mutant mosquito, removing a gene to investigate how its absence affected the insect.	这个研究小组由佛罗里达国际大学的生物学家马修·德根纳罗领导。小组通过排除法，确定了这种罪魁祸首的受体为亲离子受体 8a，简称 IR8a。这个排除法实验从 2013 年就开始了，当时德根纳罗移除了（蚊子的）一种基因，得以创造出了世界上第一只变体蚊子，为的是研究缺少这种基因对蚊子的影响	该小组由金融情报室生物学家马修·德根纳罗(Matthew DeGennaro)领导，通过 2013 年开始的消除过程，将这种负罪感受体确定为嗜离子受体 8a(简称 IR8a)，当时 Degennaro 能够制造出世界上第一只变异蚊子，去除一种基因，以调查它的缺失对昆虫的影响。

As can be seen from this table, all the names of the acronym FIU are not translated in Jinshan translation. It also mistranslated FIU into financial intelligence. In the translation of IR8ad below, it is word-to-word translation. The manual translation accurately translated FIU into Florida International University.

Human translation can avoid misunderstanding caused by improper acronyms in machine translation. For example, in the press conference of the Ministry of Foreign Affairs from 2015 to 2017, the phrase "一带一路", translated from Chinese to English by premier Li's translator Sun Ning, is different from that translated by Google.

原文：我们大力拓展对外合作，“一带一路”得到广泛的响应。特别值得一提的是，我们着力构建以合作共赢为核心的新型国际关系，正在走向结伴而不结盟的对外交往新路。

Sun Ning's translation: We made energetic efforts to expand external cooperation, and our initiative to establish a Silk Road Economic Belt and a 21st Century Maritime Silk Road won support from a lot of countries. It is particularly worth mentioning that focusing on building a new type of international relations featuring win-win. We are taking a new path of external relations characterized by partnership rather than alliance.

Google translation: We have made great efforts to expand our cooperation, and the "One Way" initiative has received a wide range of responses. Particularly worth mentioning is that we focus on building a win-win cooperation as the core of the new international relations, is out of a joint and non-aligned foreign exchange.

In the first part, the translation method of the central word "一带一路" is different for the original text. Sun Ning translated "一带一路" into "a Silk Road Economic Belt and a 21st Century Maritime Silk Road", and Google translation translated it into the "One Way". Therefore, Google translation is inaccurate since the original information cannot be delivered clearly (Zhou Yang, 2017).

Next, Mr. Sun Ning translated "得到广泛的响应" into "won support from a lot of countries" and Google translated it into "received a wide range of responses". Both expressions are feasible from grammar, but their specific meanings are different. In Sun Ning's translation, "响应" is understood as "support" the initiative to establish a Silk Road Economic Belt and a 21st Century Maritime Silk Road put forward by China. Google translation has translated it into a "response". In the light of the "New international relations based on co-operation and co-winning", the expression should be a friendly and positive relationship between China and the relevant countries.

Therefore, the translation "won support from a lot of countries" is more in line with what Foreign Minister Wang Yi wishes to emphasize. At the same time, it can be seen that "走向" in the "走向结伴而不结盟的对外交往新路" does not mean "walk". It refers to the development of Chinese foreign communication characterized by partnership rather than alliance. From the viewpoint of word use, Foreign Minister Wang Yi vividly shows the developing direction of Chinese diplomacy since both expressions are conveyed accurately, and Chinese language habits are considered.

B. Some Grammar Errors in Translation

Although machine translation based on neural network has replaced the statistical school as the main research method in the field of translation. the word error rate and grammar error rate of machine translation still exist. Here is an example of grammar error.

原文：2015 年，我们将保持进取势头，拓展全方位外交,在坚定维护国家利益的同时，不断扩大与世界各国的

共同利益。

Sun Ning’s translation: In 2015, we will continue to forge ahead and expand round diplomacy. While steadfastly safeguarding our national interests, we will work to expand the common interests we have with other countries in the world.

Google translation: In 2015, we will maintain the momentum of progress, expand all- round diplomacy, while maintaining national interests at the same time, continue to expand with the common interests of the world.

From the syntax, the subject of "拓展", "保持", and "不断扩展" are all "我们". To represent a series of actions, several verbs are connected after the same subject. The speaker had an impulse to express himself. At the same time, the compact sentence structure can also reflect how the subject describes the action, how to emphasize the personal subject itself. However, English expression habits are different from Chinese expression habits because there are so many verbs after the same subject. Sun Ning divided one original sentence into two separate sentences. Google translation missed the subject of the last clause because it wasn't broken down into sentences.

C. Lack of Natural and Fluent Sentence Structures

At present, machine translation at home and abroad mainly adopt literal translation, using simple sentences for daily communication, resulting in a lack of fluency in sentence structure. (Wang Ge, & Hua Nan, 2013).

原文	Google	human
22 岁的甘晓晨毕业于国际新闻专业，她在今年的求职过程中感到些许不轻松。	22-year-old Gan Xiaochen graduated in international journalism, and she felt a little uncomfortable during the job search process this year	Gan Xiaochen, 22, a graduate majoring in international journalism, has had a tough time finding a job this year
“她发现身边的求职者都是些海归，他们语言能力强，举止优雅，拿着一流的学位。”在四大会计师事务所的一次集体面试中，我发现 3 /8 的求职者都有海外留学经历。	“She found that the job seekers around her are all returnees. They have strong language skills, elegant manners, and hold first-rate degrees. "In a collective interview with the Big Four accounting firms, I found that 3/8 of the job applicants had overseas study experience.	She found that the job seekers around her were all overseas returnees with good language skills, polished manners, and a first-class degree. In a group interview for the Big Four accounting firm, I found three out of eight job applicants had overseas study experience.

In the first half of the sentence, there is no significant difference between machine translation and human translation. However, in machine translation, "些许不轻松" is translated into " a little uncomfortable ". "些许" corresponds to "a little" in Chinese, which is used to modify "不轻松". But it is translated into " a little uncomfortable" in Google translate, which is not accurate to its original meaning. By comparison, it is not difficult to find that machine translation is not good in dealing with the relationship between sentences. For example, Google translates "她发现身边的求职者都是些海归", into ““She found that the job seekers around were all returnees.” The meaning “海归”is ignored.

This requires the translator to intervene manually according to the original text and the machine translation. Otherwise, the translated text will be full of errors. To sum up, artificial intelligence can provide a lot of benefits to us in machine translation, there is still a big gap between artificial intelligence and artificial translation. It cannot displace human translation.

IV. CONCLUSION

According to the general characteristics of news language and its latest changes, this paper makes a comparative analysis and finds that there are still some problems in current machine translation, such as inaccurate words, rigid sentence patterns, unclear expression of specific cultural meanings, etc. Therefore, in order to further reduce translation problems, it is necessary to further study the language errors in news reports.

REFERENCES

[1] Guo Jin. (2018). Language features and translation strategies of news English. *Campus English* (21). 4-9

[2] Li Qiang. (2019). Application research of translation workshop in news translation teaching for business English majors. *China press* (06):36-37.

[3] Liu Qiyuan. (2016). Comparison and evaluation of two kinds of machine translation. *Curriculum Education Research* (28). 10(28).123-124

[4] Luo Huazhen, Pan Zhengqin, & Yi Yongzhong. (2017). Analysis of the development status and prospect of artificial intelligence translation. *Electronic World* (21):21-23.

[5] Mao Xiaoli. (2019). English translation software development and application in information environment. *Microcomputer Application*, 35(02), 120-122.

[6] Ngoc-Tien, L. , Benjamin, L. , & Laurent, B. . (2018). Automatic quality estimation for speech translation using joint ASR and MT features. *Machine Translation*, 32(4): 325-351.

[7] Qiu Yuan, Yin Rong, Wang Yamin, & Wang Xuelian. (2017). Comparative study of five translation platforms. *Translation BBS* (2).11-14.

- [8] Rob Vandenberg. (2018). How AI is changing the future of translation management. *Multilingual*, 29(08),32-35.
- [9] Tian Jiangbo. (2018). Whether machine translation can replace human translation from the perspective of translation standards. *Overseas English*, No.377(13), 132-133.
- [10] Una Zhang. (2018). Neural network after the translation of the edit case study. The Beijing second foreign languages institute (Doctoral dissertation).
- [11] Wang Dongmei. (2019). Research and countermeasures on service innovation model of cross-border e-commerce platform. *Market Weekly* (02), 79-80.
- [12] Wang Ge, & Hua Nan. (2013). The influence of semantic analysis on machine translation -- a case study of news language. *Journal of Jiangsu economic and trade college* (05), 55-57.
- [13] Wu, Yonghui, Schuster, Mike, Chen, Zhifeng.etc. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. Technical Report. <https://arxiv.org/pdf/1609.08144v1.pdf>.
- [14] Xu Xuehui. (2018). Evaluation of Chinese-English translation quality in neural network machine translation. Beijing foreign studies university.
- [15] Yin Jiao. (2017). Vocabulary and syntax error analysis in English Chinese machine translation. University of electronic science and technology of China (Doctoral dissertation).
- [16] Zhang Liang. (2019). Process reengineering of news production in the era of artificial intelligence. *Publishing wide Angle*, 333(03), 42-44.)
- [17] Zhou Yang. (2017). The CDA perspective analysis human translation with Google translation quality of political discourse translation. The Beijing foreign studies university (Doctoral dissertation).

Huiqiong Duan was born in Hunan, China in 1968. She received her PhD in English Language and Literature from Shanghai International Studies University, China in 2011. She is currently an associate professor at School of Foreign Languages, Nanchang Hangkong University, China and a visiting researcher at University of Auckland. She is a principal investigator of China National Social Science Fund project. Her research interests include cognitive linguistics and language testing.

Xinyu Hu was born in Sichuan, China in 1996. She received her Bachelor's degree in English linguistics from Nanchang Hangkong University, China in 2019. Her research interests include journalism and translation.

Yidan Gao was born in Hunan, China in 1995. She is a graduate at School of Communication, Hong Kong Baptist University. Her research interests include big data analysis.