2021 2nd International Conference on Information Science and Education (ICISE-IE) | 978-1-6654-3829-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICISE-IE53922.2021.00146

An Analysis of Errors in Google Neural Machine Translation

Yimeng Wang [1,a]
[1]Dalian University of Technology
[a]e-mail: wymssss800@mail.dlut.edu.cn

Siqi Liu [2,b]
[2]Dalian University of Technology
[b]e-mail: 20201181031@mail.dlut.edu.cn

Yingwei Ma[3,c]
[3]Dalian University of Technology
*Corresponding author: mayingwei@dlut.edu.cn

***Abstract*—Google Neural Machine Translation employs an encoder-decoder framework with an attention mechanism and it has resulted in a tremendous increase in efficiency. However, the quality of its translations is not yet comparable to that of human translations. In order to discover the specific problem of GNMT, the study analyzes the accuracy rate and types of errors in Google Neural Machine Translation by means of case studies. Based on the data collected, it is found that errors at the lexical level are most significant. The sparse data for certain terms in the Google Translate database, such as buzzwords, make undertranslation, a subcategory at the lexical level, the most significant translation error. These results proved that the users and developers may pay greater attention to the lexical level and try to match the self-learning capability of machine translation to the speed of partial vocabulary updates.**

Keywords—Google neural networks; machine translation; translation errors;

I. INTRODUCTION

Google Neural Machine Translation (GNMT) has significantly improved the quality of machine translation compared with the rule-based machine translation and statistical machine translation. However, translation errors can still be found and the gaps between machine translation and human translation are still problems need to be addressed.

Among the relevant studies, Luo Jimei and Li mei (2012) analyzed the machine translation errors based on corpus of translated automotive technical literature [1]. Zhang Yu (2020) chose an economic text as a case and analyzed the machine translation errors and post-editing strategies [2]. Yang Yuwan (2020) selected scientific books and counted the number and percentage of each error type in machine translation [3]. On the basis of the above researches, the author thinks that in specific usage contexts, many non-academic texts also need to be discussed.

In terms of classification of errors, Li Mei and Zhu Ximing (2013) refers to the "Multilingual Corpus Construction and Research in English in Chinese Universities multilingual corpus construction and research of English corpus" and classified machine translation errors into lexical, syntactic and other categories [4]. Cao Cong (2021) formulated machine translation error subcategories include redundancy, incoherence, structure misuse and so on [5]. Yan Xiaoyu (2021) got the classification that included lexical, syntactic and discourse error types [6]. This study focuses on the translation error types of GNMT in non-academic articles and considered the ratios of different types.

II. MATERIALS AND METHODS

A. Research Questions

To investigate the machine translation errors in non-academic articles, the following three questions will be used to guide this research: What is the accuracy rate of machine translation? What are the types and percentages of translation errors in machine translation? What types of translation errors are most likely to occur in machine translation?

B. Case Selection

To ensure the accuracy of the study, the author chose bilingual news from China Daily. China Daily is one of the authoritative media in China, providing domestic news, international news, commentary and other comprehensive news information. The English parts of the bilingual news are selected from foreign official media, such as Reuters and The Guardian. The Chinese parts are translated and edited by the editors. 50 news articles were selected throughout the study, with an average of 400 words per news article, covering politics, economics, culture and all aspects of life.

C. Analysis Method

At the beginning of the study, the machine translation error classification determined by Li Mei and Zhu Ximing was selected as the classification standard for this study. To accurately find all machine translation errors, each sentence obtained by GNMT was compared with the reference. In the process of conducting news case analysis, the linguistic characteristics of the news cases were considered and moderate corrections were made to the initial classification of machine translation errors. After analyzing 675 sentences in the full set of 50 cases and correcting for error types, the machine translation error classification criteria suitable for the non-academic articles in this study was gotten.

III. RESULTS AND DISCUSSION

A. Accuracy Rate of GNMT

According to Table 1, the accuracy rate of GNMT reached 22.22%. Compared with other studies on academic articles, the accuracy rate has improved. However, for the public who use machine translation or workers who need to post-edit non-

978-1-6654-3829-2/21/$31.00 ©2021 IEEE
DOI 10.1109/ICISE-IE53922.2021.00146

academic articles, the 77.78% error rate has a considerable impact on usage.

TAB.1 STATISTICS OF SENTENCES WITHOUT TRANSLATION ERRORS

Number	Type	Quantity	Proportion
1	correct	150	22.22%
2	incorrect	525	77.78%
3	Total	675	

B. Classification of Errors

As is shown in the Table 2, the translation errors are divided into lexical, syntactic, and other levels, and each type has its subcategories.

TAB.2 MACHINE TRANSLATION ERROR TYPES

Number	Errors at lexical level	Errors at the syntactic level	Other categories
1	Wrong choice of word meaning	Disorder	Symbolic mistranslation
2	Mistranslation of a reference	Mistranslation of noun phrases	
3	Lexical category mistranslation	Mistranslation of verb phrases	
4	Mistranslation of abbreviations	Mistranslation of prepositional phrases	
5	Omission	Mistranslation of clauses	
6	Undertranslation	Mistranslation of the passive voice	
7	Overtranslation		

1) Errors at lexical level

1. Wrong choice of word meaning: The machine translation incorrectly selects the Chinese meaning of the English counterpart. These errors are widely found in terms and verbs.

Example1: A funky taste in water doesn't mean you can't drink it.

GNMT:水中的时髦味道并不代表您不能喝。

REF:水有怪味不意味着不能喝。

Analysis: The word 'funky' has multiple meanings, and the machine translation doesn't select the correct one.

2.Mistranslation of a reference: A different expression is used to write a recurring event and the machine translation does not translate it into the same correct meaning. This error may also be that the pronouns (e.g., it, he) are translated directly, making the translation ambiguous.

Example2: Papua New Guinea has also been highlighted as a cause for concern…. He added that the country - which has received 140,000 vaccine doses through Australia.

GNMT:巴布亚新几内亚也引起人们的关注。他补充说，澳大利亚已经通过澳大利亚收到了 14 万剂疫苗。

REF: 巴布亚新几内亚也是令人感到担忧的一个疫区。他补充道，该国经由澳大利亚获得了 14 万剂疫苗。

Analysis: 'The country' should refer to Papua New Guinea, machine translation incorrectly translates it as 'Australia'.

3.Lexical category mistranslation: Mistranslation of word lexemes, e.g., translating nouns as verbs.

Example3: The celebratory event includes an interactive Game of Thrones spotlight page on the HBO Max streaming service, live now, ...

GNMT:庆祝活动包括 HBO Max 流媒体服务上的互动式《权力的游戏》聚光灯页面，现在直播，...

REF:庆典活动包括在 HBO Max 流媒体平台上交互式的《权游》聚焦页面直播节目，...

Analysis: The phrase 'live now' should be extended to 'which is live now', 'live' is used as an adjective in the original text and should be translated as a noun, but the machine translates it as a verb.

4.Mistranslation of abbreviations: Machine translation doesn't translate the correct meaning of an acronym or retains the English abbreviations in the Chinese translation.

Example4: What we need to do is have the DOJ also criminally prosecuting and using the penalties.

GNMT:我们需要做的是让 DOJ 也对这些事件中的任何一起进行刑事起诉和使用处罚.

REF:我们需要做的是让司法部对这些肇事者提起刑事诉讼.

Analysis: DOJ is short for 'Department of Justice'. Machine translation doesn't translate this acronym.

5.Omission: Words in the original English are not translated in the machine translation.

Example5: ...a low figure in a country of more than a 1.3 billion people.

GNMT:而在这个拥有 13 亿人口的国家中，这一数字还是很低的。

REF:这对于一个人口超 13 亿的国家来说是比较低的水平。

Analysis: 'More than' is not translated. Machine translation simply translates the original text as '1.3 billion people'.

6.Undertranslation: Machine translation does not adequately translate the text in the context, resulting in the absence of some information.

Example6: India - experiencing a second wave - recorded more than 230,000 new cases on Saturday alone.

GNMT:印度正在经历第二次浪潮，仅在星期六，就记录了超过 23 万例新病例。

REF:正在遭遇第二轮疫情的印度光是星期六的单日新增病例就超过了 23 万例。

Analysis: The word 'wave' should be translated with its derived meaning in context. Machine translation only translates the surface meaning.

Example7: An estimated 2,500 people were given fake shots.

GNMT:估计有 2,500 人被假射。

REF:估计有 2500 人被注射了假疫苗。

Analysis: The machine translation output is too abbreviated, and the context should be considered.

7.Overtranslation: Machine translation correctly translates content that does not need to be translated or incorrectly translates content that does not need to be translated.

Example8: ...introducing a four-day workweek may not necessarily encourage employees to use their time off in a way that benefits their careers or contributes to the economy.

GNMT:引入四天工作周可能并不一定会鼓励员工以有利于他们的职业生涯或为经济做出贡献的方式利用他们的休假时间。

REF:引入每周四天的工作制不一定能鼓励员工以有利于他们的职业或促进经济的方式来利用额外的休息时间。

Analysis: In English, 'may' and 'not necessarily' both means ‘uncertainty’, so when translating into Chinese, only one word is needed to be translated.

2) Errors at the syntactic level

1.Disorder: The machine translation incorrectly changes the word order of the original English text.

Example9: Carrie Symonds, who rented her wedding dress and her outfits for the G7 conference.

GNMT: Carrie Symonds 为 G7 会议租用了她的婚纱和服装。

REF:凯莉·西蒙兹曾穿着租来的衣服举行婚礼和出席七国集团峰会。

Analysis: 'For the G7 conference' is only an addition to 'outfits', but machine translation incorrectly translates this sentence in front of 'wedding dress and her outfits' and causes an ambiguity.

2. Mistranslation of noun phrases: This error refers to the mistranslation of noun combinations.

Example10: Thrones-themed barrels of wine

GNMT:以 Thrones 为主题的酒桶

REF:权游主题的桶装葡萄酒

Analysis: Incorrect translation of the noun phrase 'barrels of wine', the vocabulary is centered on 'wine' rather than 'barrels'.

3. Mistranslation of verb phrases: This error includes matching the wrong subject, splitting a verb phrase in translation and mistranslating multiple objects attached to the verb phrases.

Example11: Getting drunk on ice cream used to be the stuff of dreams

GNMT:喝冰淇淋曾经是梦想中的东西

REF:吃冰淇淋吃到醉在过去是做梦才会发生的事

Analysis: The verb phrase 'get drunk on' is mistranslated, translating only the meaning of 'drink', without combining 'dunk' and 'get' together.

4. Mistranslation of prepositional phrases: The machine translation is inaccurate in translating the meanings of prepositions such as 'with', 'from' and 'to' in different contexts.

Example12: HBO will be commemorating the 10th anniversary of Game of Thrones with the Iron Anniversary.

GNMT:HBO 将与《铁血纪念日》一起纪念《权力的游戏》十周年。

REF:HBO 将会以“铁王座庆典”来纪念《权力的游戏》开播十周年。

Analysis: 'With' in this prepositional phrase means 'in the manner of', not the meaning of 'in the company of '.

5. Mistranslation of clauses: Machine translation has difficulty in telling the structures of complex clauses, including the inability to identify the antecedent to which a relative pronoun or relative adverb refers.

Example13: HBO will surprise three couples who held Westeros-themed weddings with special anniversary gifts.

GNMT:HBO 将为三对夫妇举行以 Westeros 为主题的婚礼并提供特别的周年纪念礼物。

REF:HBO 将会给三对举行维斯特洛主题婚礼的新人送上特别的庆典惊喜礼物。

Analysis: In this clause, the antecedent is 'three couples' and the relational pronoun is 'who'. The machine translates 'HBO' as the antecedent, resulting in an error in meaning.

6. Mistranslation of the passive voice: The original English language has passive expressions, and machine translation fails to recognize the passive voice and translates it as active; Part of the passive voice in the original English is customarily translated as active, but machine translation fails to translate them into the active form.

Example14: The researchers believe that proving that a wagyu steak can be accurately 3D-printed could be a big step toward a sustainable future.

GNMT:研究人员认为，证明和牛牛排可以准确地 3D 打印可能是迈向可持续未来的一大步。

REF:研究人员认为，证实和牛排能被 3D 打印是迈向未来可持续生产的重大一步。

Analysis: The machine translation doesn't recognize the passive dynamic in the sentence and translates it as active.

3) Other categories

1.Symbolic mistranslation: Mistranslation of punctuation, brackets, units of physics, numbers, etc.

Example15: The first trailer has been released for "Spider-Man: No Way Home".

GNMT:"蜘蛛侠：无路可走"的第一部预告片已经发布。

REF:《蜘蛛侠：英雄无归》日前发布了第一支预告片。

Analysis: In Chinese, the movie title needs to be added with a book title number. However, machine translation uses quotation marks in this sentence.

C. Ranking Analysis: The most common errors in GNMT

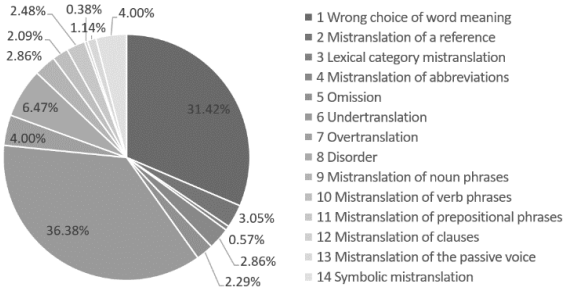

Fig.1 Proportion of machine translation error types

Figure 1 shows the percentage of error types in the 50 news articles and the data will be analyzed below. The highest percentage of errors is undertranslation, which is 36.38%. This is due to the fact that news contains a large amount of common knowledge, background information which isn't expressed in words. Some new occurrences and emerging expressions also cannot be adequately translated due to the lack of relevant data.

The second most common translation error is "wrong choice of word meaning", which is 31.42%. This is mainly due to the phenomenon of polysemy in English, and machine translation cannot find the accurate meaning of a word.

Although the error of omission accounts for only 2.29%, it is due to too little relevant data for certain words. But these words often play a key role in the sentence and influence the meaning. Therefore, omission is a problem that needs attention.

The total percentage of errors in the syntactic level is 15.43. The percentage of disorder is 6.47. This type of error happens when machine translation does not take into account the Chinese language conventions, and simply translates the original text directly. Sometimes the order of words needed to be adjusted when the text is translated, however, the machine translation still retains the original order.

The mistranslation of nouns, verbs and prepositional phrases are mainly because machine translation misjudges the way how words are combined. Such errors need to be distinguished from the mistranslations caused by errors in the selection of word meanings.

Mistranslations in the symbol category account for 4% of the errors, but the time consumed to correct these errors should be taken into consideration.

IV. CONCLUSIONS

Google Neural Machine Translation provides a tremendous aid to a person's understanding of the world and assists people to become more efficient and create more achievements. The quality of machine translation is of paramount importance.

As a case study on machine translation errors, this study only serves to reveal where the specific problems lie in the whole process of improving the quality of machine translation. This study can demonstrate the specific forms in which machine translation errors occur and promote users to identify the problems that mainly need to be addressed. Besides, this research can also help developers to find specific directions for optimizing machine translation. For users of machine translation, post-editing can be made mainly at the lexical level. The users can add some background information and increase the readability of the text. For developers of machine translation, they can focus on dealing with the error of undertranslation and take full account of the differences in terminology between Chinese and English.

This paper is devoted to help people understand the manifestations of translation errors that occur in non-academic texts, further optimize the experience of using machine translation, and promote the improvement of work efficiency.

REFERENCES

[1] Luo Jimei &Li Mei. (2012). Machine translation error analysis. Chinese Translation (05),84-89.

[2] Zhang Yu. (2020). An Analysis of the Types of English-Chinese Machine Translation Errors and Their Post-translation Editing Strategies (Master's Thesis, Beijing Foreign Studies University)

[3] Yang Yuwan. (2020). Post-translation editing of neural machine translation - An example of English-Chinese translation of Submarine Hydrodynamics. Chinese Science and Technology Translation (04),21-23+42.

[4] Li Mei & Zhu Ximing. (2013). Classification of English-Chinese machine translation errors and statistical analysis of data. Journal of Shanghai University of Technology (Social Science Edition) (03),201-207.

[5] Cao Cong. (2021). Analysis of Machine Translation Errors and Post-translation Editing Strategies for Popular Science Texts (Master's thesis, Beijing Foreign Studies University).

[6] Yan Xiaoyu. (2021). Translation report of The Oxford Handbook of Food, Politics and Society (Chapter 2) under MT+PE model (Master's thesis, Inner Mongolia University