

Research on Neural Network Post-Translation Editing——Take Goggle Translation of *To Live* as an Example

Zhang Yuanyi

Southwest Petroleum University, Chengdu, China

Abstract: From the perspective of category conversion theory, this paper studies the post-translation editing of neural network translation by taking the English version of *To Live* translated by Google as an example. The study found that Goggle can basically complete the language category conversion, but it applies the word-to-word translation method to most of the non-linguistic category conversion, making the translation confusing. Therefore, in post-editing, the translator should consider the cultural background of the target language and select the category prototype in line with the cognition of the target readers; in the process of vacancy category conversion, the translator should select the translation methods such as the conversion of similar categories to improve the reading experience of the target readers.

Keywords: neural network machine translation; post-editing; category conversion; *To Live*

1. Introduction

In November 2016, Google officially launched Neural Machine Translation (NMT). It is said that NMT can refer to the context of the whole paragraph and sentence when translation, and constantly "learn" so as to improve translation and make it more natural and reasonable. Post-translation editing of machine translation refers to "the process of improving the quality and applicability of translation through manual review and editing"⁰. If translators use NMT for basic translation, and then do the post-translation editing themselves, they would get higher quality translation in a shorter time.

According to the 2016 Global Language Service Market Report by Common Sense Advisory, machine translation and post-editing accounted for 39.4% of the entire language service market in 2016, and 25.39% of language service providers provided post-editing services. And they believe that this service has a bright future⁰. Machine translation + post-translation editing can help translators cope with large volumes of text, accelerate turnaround times, achieve greater accuracy, and increase efficiency as mentioned in *Why Machine Translation with post-editing Is the Future of Translation*, published by Translate. One in June 2020⁰.

At present, there are papers about the explanations on post-translation editing in China, such as Cui Qilang's, Zhao Tao's⁰, but there are few papers combined with cases analysis; there are studies on post-translation editing of practical texts, such as Feng Quangong's and Yang Yuwan's⁰; but there are few studies on post-translation editing of literary texts. As a novel, *To Live* has simple language, strong narrative and many words with high context and cultural dependency. The study of post-translation editing is a combination of case study and literary text study. Based on the perspective of category conversion, this paper takes *To Live* translated by Google as an example to study the post-translation editing of neural machine translation.

2. Overview of Category Conversion

Categorization is a psychological process in which human beings recognize, classify and understand something according to a specific purpose⁰. Catford first mentioned the conversion of language categories in the *A Linguistic Theory of Translation*⁰. Taylor added that categories can be roughly divided into two categories: linguistic category and non-linguistic category⁰. Linguistic category refers to the category in natural language, such as part of speech category conversion, syntactic category conversion, semantic category conversion, stylistic category conversion, pragmatic category conversion, etc.; non-linguistic category refers to things existing in the external world, such as animals and plants, appliances,

history and culture, etc. Wen Xu believes that there are three main forms of non-linguistic category conversion, namely, equivalent category conversion, Mismatch category conversion and vacancy category conversion⁰. Equivalent category can be directly converted; the methods of category Mismatch conversion can be divided into the conversion of category prototype and the conversion of category level; The methods of vacancy category conversion can be divided into the conversion of similar category and the conversion of category transplantation. Combining Taylor and Wen Xu on the classification and conversion of categories, figure 1 can be obtained.

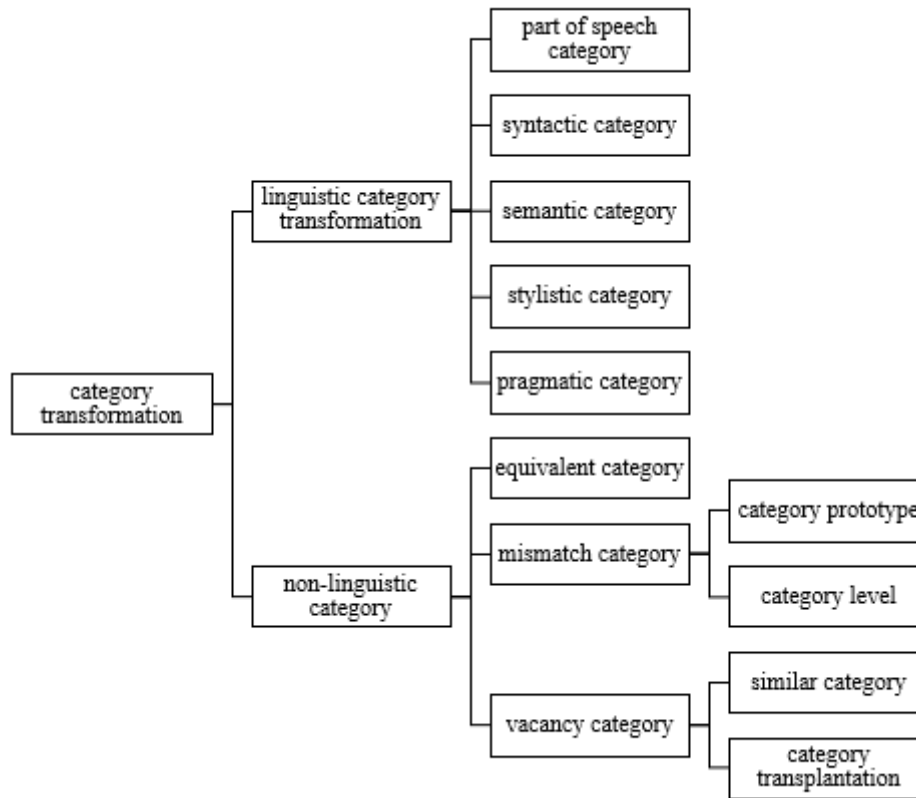


Figure 1: The classification of category conversion

Equivalent category conversion in non-linguistic category conversion refers to the existence of corresponding category members in the target language, such as "苹果" in Chinese corresponding to "apple" in English.

Mismatch category refers to the non-correspondence between the members of the source language category and the members of the target language category. It is divided into the non-correspondence of category prototype and non-correspondence of category level. Category prototype refers to the typical members of the category. For example, in the category of "mammals", that the lion, cattle and sheep are typical members, but the whale is not typical member. Rosch divided the category system into two dimensions: horizontal and vertical⁰. Horizontal category conversion refers to replacing member A in the source text with member B in the target text. A and B are within the same category. For example, in the color category, "白发" in Chinese is converted to "gray hair" in English. The vertical level is a vertical affiliation, which is divided into up level category (such as animals), basic level category (such as tigers) and lower level category (such as Amur Tigers). In vertical level, the more up the category is, the more abstract it becomes; the more down the category is, the more specific it becomes. Moreover, the basic level category has cultural and regional characteristics. Sometimes, for people of one language, it is a basic category, but for people of another language, it may not be a basic category⁰.

Vacancy category means that the category in the source language cannot find the corresponding category in the target language; vacancy category is common in words that are highly depend on the contexts and culture⁰. Vacancy category conversion methods are divided into similar category conversion and category transplantation conversion. Similar category conversion means that although the source language categories do not exist in the target language category, there are similar categories in the target language, which can roughly convey the information that to be conveyed by the source language, such

as "阎王" in Chinese and "Death" in English. Category transplantation refers to the direct transplanting of the source language category to the target language category. In this way, the target language readers have the opportunity to know the cultural customs of the source language country. For example, "Athena", the God of wisdom and war in Western culture, is transliterated into "雅典娜".

The study of machine translation post editing from the perspective of category conversion can help analyze and improve the machine translation of literary works and make the translation more acceptable to the target readers. Especially for the local literary works with many high context dependent words such as Chinese work *To Live*, category conversion can provide new ideas for translators to improve the work of machine translation.

3. Introduction to *To Live*

To Live is a local literary work created by Yu Hua in 1992, with about 90000 words. In the book, the protagonist Fugui was born in a landlord family with a wealth of money. However, due to his lust and gambling, his wealth dissipated. The social changes made the situation worse. Finally, all his family were died, leaving him alone.

Zhou Lingshun believes that local literary works have "local characteristics". The works take "local language" as the carrier⁰. *To Live* is a work of Chinese local literature. There are many unique categories of Chinese local literature, such as "风水先生", "花花绿绿", and so on. These words have the local characteristics, and they reflect the local customs to a certain extent, and need to be understood in a specific language environment. Therefore, they are a certain challenge to machine translation, and the translator has the necessity of post-translational editing. At present (January 2022), there are 142 papers study the translation of *To Live* on CNKI, a famous paper platform in China. These papers study it mainly from the perspective of translation variation theory, such as Xu Qiong's and Wei Fang's⁰ and aesthetics, such as Zhu Zhenwu's⁰. There is no paper studying it from the perspective of category conversion. However, one of the most basic aspects of cognition is the ability to divide categories⁰. Translation research from the perspective of categorization can help translators jump out of the traditional translation framework and create better translations based on the cognition of target readers. In view of this, based on the perspective of category conversion, this paper studies the post-translational editing of Goggle translation of *To Live*, which is conducive to promoting the further development of machine translation and enriching the research perspective of *To Live*.

4. Post-Translation Editing of Goggle Translation's *To Live* from the Perspective of Categorization

Categorization is a high-level cognitive activity of human beings to classify all things in the world. On this basis, human beings have the ability to form concepts, and language symbols have their meanings⁰. Categorization is one of the most important cognitive abilities of human beings⁰. Translation is the mutual conversion of two languages. Post-translation editing from the perspective of category conversion can help the translator make the machine translation more in line with the cognition of the target language readers and make the translation more accurate, smooth, and natural. Based on the perspective of category conversion, the following takes the English version of *To Live* translated by Goggle as an example to conduct post-translational editing research.

4.1 Linguistic Category Conversion

Linguistic category conversion refers to the conversion of language itself, such as part of speech category conversion, tense category conversion, etc. (see Figure 1 for details). Linguistic category conversion is very common in translation, such as example 1.

Example 1:

ST: 她张着嘴向我笑。 (P.64)

GT: She smiled at me with her mouth open.

Example 1 is the conversion of part of speech. In the source text, there are two verb phrases, "张着嘴" and "笑"; in the translation, there is only one verb "smile". The Goggle converted "张着嘴" to "with her mouth open", a prepositional phrase, because in English only one verb is allowed in one sentence. By conversion, Goggle made the translated version comply with the rules of English grammar.

The conversion of linguistic category does not take culture factors into account, and Goggle can handle it well. Since Goggle translation (GT) is correct, there is no need to carry out post-translational editing (PE). Therefore, the conversion of linguistic category would not be explored too much here.

4.2 Non-Linguistic Category Conversion

Non-linguistic category includes material categories (such as animals and plants, appliances, etc.) and abstract categories (such as history and culture, etc.), which is the human's cognitive classification of the external world. Due to different geographical and cultural backgrounds, Chinese and people in other countries have different experience. The philosophical basis of categorization theory is empirical realism philosophy, which means that the categories aroused by Chinese and non-Chinese readers when they see the same object may also be different, that's when category conversion is needed. The following explores the non-linguistic category conversion from three aspects: equivalent category conversion, Mismatch category conversion and vacancy category conversion.

4.2.1 Equivalent Category Conversion

Equivalent category refers to the existence of corresponding category members in the target language. The source language and the target language are one-to-one correspondence. If the two categories are equivalent, translator can convert them directly, such as example 2.

Example 2:

ST: 像一把钝刀子 (P.24)

GT: Like a blunt knife

In example 2, the ST and GT are completely equivalent. "钝" is equivalent to "blunt", and "刀子" is equivalent to "knife".

Equivalent category conversion is very simple, and it is mastered by most people. It is not the key and difficult point of this research.

4.2.2 Mismatch Category Conversion

Mismatch category conversion often involves the treatment of cultural differences. In translation, translators should consider the target language readers' cognition of the category, and then carry out category conversion. Although neural network translation has made great progress compared with previous machine translation, there still exists a problem of word-for-word translation, which needs the translator to carry out post-translational editing.

Mismatch category conversion methods are divided into category prototype conversion and category level conversion. The following carries out the research on post-translational editing from these two aspects.

4.2.2.1 Category Prototype Conversion

Example 3:

ST	妓女...说我心肠黑。 (P.12)
GT	The whore said I had a dark heart.
PE	The whore said I had a cold heart.

In Chinese, "黑心" refers to a person who is cold or inhuman, not a person who has a black heart. The phrase "a dark heart" in GT is a word-for-word translation. If want to express "indifference", "黑心" and "冷血" in Chinese can be used. But in English, "black" does not have the same meaning, it is just a color. The English word "cold" includes the meaning of "indifference", so the translation of "心肠黑" can be translated as "cold heart".

Example 4:

ST	于是村里人就知道那个会讲荤故事会唱酸曲的人又来了。 (P.3)
GT	So the people in the village knew that the person who could tell stories about meat and sing sour music was coming again.
PE	Therefore, the people in the village knew that the person who could tell blue/dirty stories and sing sad music was coming again.

The Chinese category "荤" generally means "meat" as opposed to "vegetable". With the development of social culture, "荤" has the meaning of "pornography" in some contexts. In this sentence, "荤故事" means "stories about porn". However, in Western culture, the category of "meat" does not mean "pornography", so it is inappropriate for Google to translate it into "stories about meat". "Blue/dirty" is often used to describe "pornography" in the West. It is suggested to replace "meat stories" with "blue/dirty stories".

Similarly, "唱酸曲" is not "sing sour music". In the category of "酸", the meaning of "sad" is included. But in the category of "sour", there is no such meaning. "Sour" is just a kind of taste. It is suggested to change "sing sour music" to "sing sad music".

As can be seen from the above examples, there are three steps in the process of category prototype conversion: 1) source language category decoding. Generally, a category often contains multiple meanings. It is necessary to know what these meanings are; 2) word meaning selection. After category decoding, according to the specific context of the source text, select the word meaning that is most consistent with the source text; 3) target language category selection. If there is equivalent category in the target language, the category conversion can be carried out directly; if the category is not equivalent, then consider the cultural background of the target language and choose the category that conforms to the prototype cognition of the target language readers.

4.2.2.2 Category level conversion

Category level can be divided into horizontal category level and vertical category level. The first one is the horizontal level conversion, as shown in example 5.

Example 5:

ST	他脸上青一块紫一块的。 (P.54)
GT	His face was blue and purple.
PE	His face was black and blue.

In the category of color, there are some differences between Chinese and English. For example, if we want say that a person's skin injury is serious, Chinese use "青"(the color which is approximately corresponding to "blue" in English), and "紫"(corresponding to "purple" in English), such as "青一块紫一块", "鼻青脸肿" and so on. However, in English, "blue and purple" does not use to describe injury. The Chinese phrase "青一块紫一块" is equivalent to English phrase "black and blue".

Another is the conversion of vertical levels, such as example 6.

Example 6:

ST	我女人家珍当然知道我在城里这些花花绿绿的事。 (P.12)
GT	Of course, my woman Jiazhen knows all these colorful things about me in the city.
PE	Of course, my wife Jiazhen knows all these colorful things about me in the city.

The Chinese category of "女人" is an up level category, including 1) "girlfriend or wife"; 2) "female". In this sentence, the category of "女人" refers to "wife". The English category of "woman" is only a classification of gender, so it is inappropriate to directly translate "女人" into "woman". It should be translated to "wife".

But when dealing with "花花绿绿", Goggle handled it properly. When Chinese see "花花绿绿", they think of flower and grass, the color of red and green. In the *To Live*, "花花绿绿" means "my" colorful life, including gambling, visiting brothels and teasing "my" father-in-law. "花花绿绿" belongs to the category of basic level. However, in English, "flower" and "grass" can not reflect the colorful life of "me". According to Collins Advanced English Chinese Dictionary, the basic meaning of "colorful" is "having bright colors or a lot of different colors", which is an up category. One of its extended meanings is "If someone has had a colorful past or a colorful career, they have been involved in exercising but often slowly shooting things.", which is consistent with the information to be conveyed by the original text. Goggle did it well.

Color category conversion in horizontal category conversion is typical and special, because color not only refers to color, but also reflects the common basis of human neurophysiology and different social and cultural connotation. Therefore, when translating, translators need to consider the cognition of target readers' color category. In vertical level, the upper lever the category belongs, the more abstract its meaning becomes. When it is necessary to change category level, first the translator should consider that,

in target readers' cognition, which level the category belongs, and then select the appropriate category level for conversion.

4.3 Vacancy Category Conversion

Vacancy category generally depends much on culture background, only appearing in one its own culture. Since no equivalent category can be found in the target language, it is a big challenge for machine translation. Usually, machine translation will use Pinyin or similar categories to carry out conversion, but the results are often unsatisfactory. At this time, the translator needs to carry out post-translational editing.

The methods of vacancy category conversion are divided into similar category conversion and category transplantation conversion.

4.3.1 Similar Category Conversion

Example 7:

ST	我是我们徐家的败家子.用我爹的话说.我是他的孽子。 (P.7)
GT	I am the prodigal son of our Xu family. In my father's words, I am his sinner.
PE	I am the prodigal son of our Xu family. In my dad's words, I am a bastard.

In Chinese dictionary, one explanation of the word "孽" is "孽障", that is, "the words of elders scolding their feckless, unfilial descendants". "孽子" means "by-blow", but in the text, it refers to "孽障". The explanation of "sinner" in Longman Contemporary English Dictionary is "someone who has sinned by not obeying God's laws". In the *To Live*, "I" went gambling, went to brothels and beat wife, which really violated God's laws to a certain extent. However, there is no religious element in the original text, so it is not very appropriate to replace "孽子" with "sinner". Oxford Dictionary explains "bastard" as "used to inspect sb, especially a man, who has been rude, unpleasant or cruel". It can basically convey the Chinese word "孽子", and there is no additional religious element. Therefore, it is more appropriate to use to replace "孽子". Unfortunately, neither "sinner" nor "bastard" expressed "elders scolding their descendants". However, translation is untranslatable to a certain extent, so translators only need to translate its main meaning. In addition, "爹" is an informal expression, and it should also be translated into "dad", not "father".

4.3.2 Category Transplantation and Conversion

Example 8:

ST	队长陪着风水先生来到了我家门口。 (P.83)
GT	The captain accompanied Mr. Feng Shui to my door.
PE	The captain accompanied Fengshui master, who could tell the future by surveying the local geography, to my door.

"风水先生" refers to an expert who predicts good or bad luck by observing geological terrain, not a man surnamed "Feng Shui", so GT of "Mr. Feng Shui" is inappropriate. "风水" is a word with high context and cultural dependence. In Western culture, there is no similar category. It can only be transplanted. In translation, this word can be slightly explained, such as the modified translation in example 8, so as to minimize the difficulty of the target readers in understanding the translation.

5. Conclusions

From the perspective of category conversion, this paper takes Google translation of *To Live* as an example to study the post-translational editing of neural network. It is found that Google's neural network translation system can basically complete the language category conversion and equivalent category conversion, but most of the high context and cultural dependent words in *To Live* are translated by word to word or Pinyin. Therefore, this paper proposes that translators can use Google to assist translation, but when encountering category prototype conversion, translators should first decode the category in source text, then select the appropriate word meaning, and finally select the category that conforms to the cognition of the target readers; horizontal category conversion is generally the conversion of color category. The translator needs to understand the target readers' cognition of color category; when need to carry out the conversion of vertical category level, translator should first know the category in source language corresponds to which level in target language, and then select the appropriate category level for conversion; when the category is vacant in target language, first find whether there is similar category. If

so, carry out similar category conversion, if not, the category can be transplanted with annotation to help target readers to understand the text.

References

- [1] DePalma, D.A. *Post-edited machine translation define* [EB/OL]. <http://www.commonseadvisory.com/AbstractView.aspx?Arti-cleID=5499>. (2013-04-30) [2014-02-21]
- [2] Feng Quangong, Liu Ming (2018). *A three-dimensional model of post-translation editing Ability* [J]. *Foreign Language Studies*, no.03, pp.55-61.
- [3] <https://www.translate.one/2020/06/16/why-machine-translation-with-post-editing-is-the-future-of-translations/>
- [4] Cui Qiliang (2014). *On post-translation editing in machine translation* [J]. *China Translation*, vol.35, no.06, pp.68-73.
- [5] Zhao Tao(2021). *Current situation and problems of post-translation editing in machine translation* [J]. *Foreign Language Teaching*, vol.42, no.04, pp.100-104.
- [6] Feng Quangong, Li Jiawei(2016). *Research on post-translation editing mode of news translation* [J]. *Audio-visual Teaching of Foreign Language*, no.06, pp.74-79.
- [7] Yang Yuwan(2020). *Post-translation editing of neural machine translation: A case study of submarine hydrodynamics* [J]. *China Science and Technology Translation*, vol.33, no.04, pp.21-23+42.
- [8] Ungerer, F., Schmid, H.J(1996). *An introduction to cognitive linguistics*[M]. London: Longman.
- [9] Catford, J.C(1965). *A Linguistic theory of translation*[M]. London Oxford University Press.
- [10] Taylor, J(1995). *Linguistic categorization: Prototypes in linguistic theory*[M]. Oxford: Clarendon Press.
- [11] Wen Xu, Xiao Kairong(2019). *Cognitive translation* [M]. Beijing: Peking University Press.
- [12] Rosch, E. 1978. *Principles of categories*. In E. Rosch, &B. Lloyd (eds.) *Cognition and categorization*. Hillsdale, N J: Erlbaum.
- [13] Li Fuyin. *Introduction to cognitive linguistics* [M]. Beijing: Peking University Press, 2008:111.
- [14] Yu Ping(2020). *Cognitive research on category transformation in translation* [D]. Southwest University.
- [15] Zhou Lingshun(2016). *A Study on the translation of local language and the criticism* [J]. *Foreign Language Studies*, vol.33, no.04, pp.77-82.
- [16] Xu Qiong (2015). *Translation variation of Yu Hua's novel To Live* [J]. *Chinese Construction*, no.33, pp.73-74.
- [17] Wei Fang(2015). *On the English translation strategies of Yu Hua's To Live from the perspective of translation variation* [J]. *Chinese Construction*, no.27, pp.45-46.
- [18] Zhu Zhenwu, Yang Leipeng(2016). *Michael Berry's aesthetics and cultural responsibility of translation—A case study of the English translation of To Live* [J]. *Foreign Language*, vol.32, no.03, pp.89-94.
- [19] Jackendoff, R. S(1985). *Semantics and cognition*. Cambridge, MA: MIT Press.
- [20] Shao Weiyong (2017). *Abstraction and concretization -- cognitive comparison of English and Chinese expressions and category conversion in translation* [J]. *Journal of East China University of Science and Technology (Social Science Edition)*, vol.32, no.03, pp.102-109.
- [21] Wen Xu(2014). *Cognitive basis of language* [M]. Beijing: Science Press.