ISSN Print: 2707-2622 Online: 2707-2630
Linguistics− 2025, 7(2)

https://doi.org/10.35534/lin.0702012

A Comparative Study of Human and Online Machine Translation of English Passive Sentences

Linbing Wang

East China University of Science and Technology, Shanghai

Abstract: In recent years, with the continuous development of computer technology, machine translation has been more and more applied to English-Chinese translation, and its important role in the translation field has become increasingly prominent. However, due to the great differences in expression between Chinese and English, many problems have arisen in machine translation between English and Chinese. Previous related studies have only concluded that machine translation is not as good as human translation, but failed to specifically point out the problems that arise when translating certain types of English sentences using machine translation, making it difficult to provide specific improvement suggestions for machine translation. Therefore, this paper takes English passive sentences, a hot topic in the translation field, as a concrete example, compares the translations of human translation and machine translation respectively, explores the differences between them, and analyzes the reasons for the differences, aiming to sum up the problems that occur in machine translation of English passive sentences and putting forward concrete suggestions for improvement.

Keywords: English passive sentences; human translation; machine translation; comparative study

1 Introduction

In the 1930s, French scientist G.B. Artsouni proposed the idea of using machines for translation. Since then, machine translation has made technological advancements from rule-based to statistical-based and then to neural network-based approaches, and humans are now able to accurately translate foreign works. However, in the field of machine translation, "technology comes first", and developers often overlook guidance from linguistics and translation theories due to their specialized technical academic backgrounds. Related articles questioning whether machine translation can replace human translation, whether machine translation can translate cultural connotations, and the advantages and disadvantages of machine translation have emerged one after another. It is therefore clear that exploring the differences between human translation and machine translation is a hot topic in current research. However, previous studies have mostly focused on the differences between machines and humans, only summarizing some general

Corresponding author: Wang Linbing, East China University of Science and Technology.
Article Citation: Wang, L. B. (2025). A Comparative Study of Human and Online Machine Translation of English Passive Sentences. *Linguistics, 7*(2), 114-128.

problems and failing to identify specific problems that arise when translating certain types of sentences using machine translation. The topic was too broad and lacked practicality.

Passive voice is a relatively common grammatical phenomenon in English that plays an important role in grammar, pragmatics, and discourse coherence. However, due to its semantic relationship being the opposite of active voice sentence structure and its extremely complex logical relationship and syntax structure, it poses great challenges for machine translation. Therefore, taking English passive sentences as a starting point to explore the differences between human translation and machine translation can not only provide practical and effective suggestions for machine translation of specific English sentences, but also provide new ideas and vitality for the research topic of English passive voice.

2 Literature Review

2.1 Research of Machine translation

Machine translation simulates the cognitive calculation of the human brain in the process of translation in a computational way (Gao & Zhao, 2020). Compared with human translation, the biggest advantage of machine translation lies in its fast processing speed, and the biggest weakness lies in the difficulty of guaranteeing the translation quality. The reason is that it is not easy for a machine to simulate the complicated translation process of the human brain. Therefore, machine translation research has always been pursuing the improvement of translation quality. Many studies focus on the quality evaluation of machine translation, including the classification of machine translation errors. The research results in recent years show that common errors in machine translation include mistranslation, over-translation, omission, non-translation, and so on. Another kind of research takes human translation as a reference version and finds the shortcomings of machine translation by comparing human-machine translation.

The passive sentence is a common sentence pattern in English and Chinese, and its grammatical, textual, and pragmatic functions are important. However, English and Chinese belong to different language families, and the subject-predicate relationship is the basic sentence structure of English; Chinese, on the other hand, is parataxis-oriented and semantic-based, with a topic-explanation sentence structure (Wang & Wei, 2002). Moreover, English and Chinese users have different cognitive styles, so English and Chinese passive sentences are obviously asymmetrical. In addition, the inverted order of English passive sentences, "agent-patient," poses significant challenges for machine translation as the semantic logical relationship is opposite to the syntactic order (Li & Zhu, 2013). Therefore, translating English passive sentences into Chinese has always been a difficult point, and it is also crucial for breakthroughs in machine translation at the syntactic level, which has attracted the attention of some scholars. Wang (1984) summarized some rules of formal passive sentences in machine translation from the perspectives of grammatical structure and semantics of verbs. Bai and Zhan (2003) believed that when translating English passive sentences with machine translation, a large number of non-"被" words should be produced. However, most of the above studies start from the perspective of language rules, with little analysis of the language structural characteristics presented in the overall translation.

2.2 Research of Passive Voice

A great amount of research work has been done on passive voice both at home and abroad. According to the representative of traditional grammar, Quirk et al (1985), voice is an important grammar phenomenon which functions to describe either from the active or from either actively or passively, keeping the facts conveyed the same. Besides, voice is quite a common grammatical phenomenon, and the voice system of nearly all grammatical phenomena, and the voice system of nearly all languages in the world consists of active and passive voice. Further, he classified passive voice into several types. However, both his definition and his classifications of passive voice suffered from severe criticism in later days. It is not absolutely the truth that active-passive alteration never causes a change in the facts reported, and he did not incorporate semi-passives and pseudo-passives into the voice system.

Formalists and functionalists both have made great contributions to the study of passive voice, but both have their limitations. It is the study made by the cognitive approach that sheds light on our exploration of passives nowadays. Professor Wang Zhijun proposes that voice reflects essentially the language user's perspective in an event from the cognitive aspect. He claims that any event can be described either from the causative event view (cause-become-state), or the inchoative event view (become-state), or the stative event view (state). The causative view of an event produces active voice, the inchoative view produces medio-passive, and the stative view produces stative voice. The medio-passive, which construes a causative event from the inchoative view, drops a hint of process. The stative passive, which construes a causative event as a state, only designates the resultant state of the verbal process, without any implication of an external cause. Professor Wang classifies these two kinds of passives——medio-passives and stative passives into non-prototypical passives. In his eyes, prototypical passives make the affected entity in the causative event the subject; with no change in the process they describe, compared with the corresponding actives. Prototypical passives include agentless passives and agentful passives. This paper will adopt Professor Wang's classification of passive sentences to choose research corpus.

3 Research Design

3.1 Research Purpose

The purpose of this research is to explore the differences between human translation and machine translation of English passive sentences, analyze the reasons for the differences, and summarize the problems that arise in machine translation of English passive sentences, so as to provide reference and suggestions for future machine translation of English passive sentences.

3.2 Research Questions and Hypothesis

The research attempts to answer the following questions:

Question 1: Is there a significant difference between machine translation and human translation in translating English passive sentences?

Question 2: Compared to human translation, what progress and shortcomings does machine translation have?

Question 3: What are the reasons for these shortcomings?

It is hypothesized that there exist differences between human translation and machine translation, and humans do better than machines in translation.

3.3 Research Subjects

In the part of human translation, an MTI student majoring in translation from Shanghai International Studies University was selected, who has passed TEM 4 and TEM 8. After entering school, she received professional training in translation theory and practice, so she is qualified to do the corresponding English translation.

In the part of machine translation, the author searched for a lot of translation software, and finally chose three kinds of translation software with high accuracy and more use, namely, ERNIE Bot (文心一言), Chatgpt 3.5, and Deepl.

3.4 Research Corpus

Professor Wang's classification of passive sentences is adopted to choose research corpus. According to professor Wang, English passive sentences are divided into two categories, prototypical passives and non-prototypical passives. Prototypical passives include agentful passives and agentless passives. Non-prototypical passives consist of stative passives and medio-passives. That is to say, English passive sentences are divided into four categories. According to the above classification criteria, four English passive sentences are randomly selected from different corpora. Finally, 16 English passive sentences are selected as the final research corpus. See the Appendix.

In terms of the research corpus of human translation, to ensure the accuracy of the research, the author has randomly mixed up the order of various types of English passive sentences selected from the corpus and did not inform the subject that she needed to translate English passive sentences. The subject was allowed to consult a dictionary, and there was no set time limit for the translation, with the aim of producing the best possible translation within the subject's ability range.

In terms of the research corpus of machine translation, the selected 16 English passive sentences were input into three kinds of translation software, respectively, and the results of machine translation were obtained.

3.5 Statistical Software

For the research, the selected statistical software not only needs to be able to visualize and present a large amount of data, but also needs to be user-friendly and easy to operate in order to simplify the process of obtaining statistical results. SPSS software, developed by IBM, is widely used in statistics due to its simplicity and ease of use. With basic computer operation skills and some learning, users can easily utilize its powerful features to meet their analytical needs, and the output results are also straightforward and comprehensible. Excel software, on the other hand, is the most commonly used data processing software, with simple operations such as calculating averages and sums. Therefore, SPSS and Excel are selected for this research.

3.6 Research Procedures

3.6.1 The Scoring Criteria

In order to formulate objective scoring criteria and obtain valid research data, the author referred to the scoring standards of the national Translation Qualification Examination and TEM-8 sentence translation. The scoring standard of the national Translation Qualification Examination adopts the deduction system, which pays attention to the understanding of the original text. Even if the translation is stylistically beautiful, it is incorrect if it deviates from the original text. The scoring of sentence translation in TEM-8 adopts the mode of grading. According to the scoring standards of the above two kinds of exams, the author has formulated the scoring criteria suitable for this research. As shown in Table 1.

Table 1 Scoring Criteria

Criteria \ Score	5	4	3	2	0
Faithfulness (60%)	All the information from the original text is conveyed, and the tone and style are consistent with the original.	Except for a few omissions, all the important information from the original text is conveyed, and the tone and style are consistent with the original.	There are minor misunderstandings or a few mistranslations, but the main idea is consistent with the original.	There are major errors or mistranslations, some information is mixed up, but the overall meaning is generally conveyed.	There are many mistranslations and omissions, failing to convey the main idea of the original text.
Fluency (40%)	The sentence structure is handled appropriately, word choice is appropriate, paragraphs and sentences are coherent	Word choice is relatively correct and appropriate, sentence organization and arrangement conform to Chinese norms.	It is constrained by English sentence structure, the text is not smooth enough, but there are no major errors in word choice or sentence structure.	Sentences are incoherent, the text is obscure, and there are a few major errors in word choice and sentence structure.	Word choice is inappropriate, the text is incoherent, and the language does not conform to Chinese norms.

As shown in the table above, the scoring criteria consist of two parts: faithfulness (60%) and fluency (40%). Among the three principles of Skopos theory, the most important is the purpose principle, so all translations should first meet the purpose principle. Therefore, if a translation has serious mistranslation issues and fails to convey the original spirit, the sentence will receive a score of 0. Both faithfulness and fluency are divided into five levels, corresponding to 5 points, 4 points, 3 points, 2 points, and 0 points. The reason for not using a percentage system is for the convenience of data statistics and to simplify the calculation process. Additionally, specific scores are assigned directly without setting score ranges to reduce errors and avoid different evaluations of translations due to subjective errors or environmental influences.

3.6.2 Data Collection

For this research, senior students majoring in English from a certain university were selected as subjects. The original English text, machine-translated text, and human-translated text were compiled into a questionnaire and distributed to the subjects for evaluation. The subjects were required to score the translated texts based on the scoring criteria, their knowledge, experience, and personal feelings. The passive voice in the original English text was highlighted in bold to draw the subjects' attention to the translation of passive voice. The questionnaire itself did not contain any explicit instructions or hints regarding the presence of machine-translated texts to avoid potential bias and maintain the objectivity of the survey. Finally, Excel was used to collect the data.

3.6.3 Data Analysis

A total of 27 valid questionnaires were collected for this research. Excel was used to statistically analyze and calculate the average scores of 16 sentences in human translation and machine translation. In order to verify whether the hypothesis of this research is valid, we need to test the samples. The samples in this research consist of translations from three different machine translations and human translations, totaling four independent samples. Since the sample size is less than 30, if the data corresponds to the normal distribution and the homogeneity (equality) of variance, it is suitable for Analysis of Variance. Therefore, the first step is to determine whether the samples follow a normal distribution. A normality test on the samples of human translation and machine translations was conducted through SPSS, and the results are shown in Table 2.

Table 2 Normality Test

Type		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Score	ERNIE Bot	0.157	16	0.200*	0.918	16	0.154
	Chatgpt	0.23	16	0.023	0.89	16	0.056
	DeepL	0.182	16	0.163	0.928	16	0.226
	Human	0.156	16	0.200*	0.906	16	0.101

Note: *. This is the lower bound of the true significance level.
a. Lilliefors Significance Correction

Since the data size is between 3 and 50, we need to refer to the results in the right column of the Shapiro-Wilk test in the table. According to the criteria for normality testing, when the significance value in Shapiro-Wilk is greater than 0.05, the sample corresponds to a normal distribution. Based on the research data, the significance values for the four samples are 0.154, 0.056, 0.226, and 0.101, all of which are greater than 0.05. Therefore, the data from this research follows a normal distribution.

A Levene 's test of equality or error variances on the samples of human translation and machine translations was conducted through SPSS, and the results are shown in Table 3.

Table 3 Levene's Test for Equality of Error Variances[a]

Dependent Variable: Score

F	df1	df2	Sig.
2.74	3	60	0.051

Note: Tests the null hypothesis that the error variance of the dependent variable is equal across all groups.
a. Design: Intercept + Type

As shown in Table 3, the significance value (0.051) is greater than 0.05, which means the data corresponds to the equality of variance. So the data of this research is suitable for Analysis of Variance.

Next, a single-factor analysis of variance is conducted. The scores of human translation and machine translation were grouped together as the dependent variable and represented by “Score”. A grouping variable “Type” was used as the independent variable, including four levels: 1 represents “ERNIE Bot”, 2 represents “Chatgpt”, 3 represents “Deepl”, and 4 represents “Human”. By following the Analysis of Variance procedure in SPSS, the following research analysis data was obtained.

Table 4 Analysis of Variance

Between-Subjects Factors(a)			
		Value Label	N
Type	1	ERNIE Bot	16
	2	Chatgpt	16
	3	DeepL	16
	4	Human	16

Dependent Variable: Score

Descriptive Statistics(b)			
Type	**Mean**	**Std. Deviation**	**N**
ERNIE Bot	4.0806	0.29313	16
Chatgpt	4.1488	0.18191	16
DeepL	4.0994	0.31915	16
Human	4.7963	0.17212	16
Total	4.2813	0.3898	64

From the data in Table 4, it can be observed that the average value of human translation (4.7963) is higher than the average values of the other three types of machine translations (4.0806, 4.1488, 4.0994), indicating that the acceptability of human translation is higher than that of machine translations. In terms of standard deviation, Deepl has the highest standard deviation (0.33115), indicating a higher level of variability in the scores of Deepl's translation. This means that some of its translations have relatively high scores while others have low scores, showing a larger fluctuation range. In other words, the translation quality of Deepl is unstable. ERNIE Bot follows with the second-highest standard deviation (0.29313). Human (0.17212) and ChatGPT(0.18191) have smaller standard deviations, indicating a more balanced distribution of scores for these two, resulting in a more stable quality of translation.

Table 5 Paired Sample t-Test

Dependent Variable: Score

(I) Type	(J) Type	Mean Difference (I - J)	Std. Error	Sig.b	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
ERNIE Bot	Chatgpt	-0.068	0.09	1	-0.313	0.177
	DeepL	-0.019	0.09	1	-0.264	0.226
	Human	-0.716*	0.09	0	-0.961	-0.47
Chatgpt	ERNIE Bot	0.068	0.09	1	-0.177	0.313
	DeepL	0.049	0.09	1	-0.196	0.295
	Human	-0.648*	0.09	0	-0.893	-0.402
DeepL	ERNIE Bot	0.019	0.09	1	-0.226	0.264
	Chatgpt	-0.049	0.09	1	-0.295	0.196
	Human	-0.697*	0.09	0	-0.942	-0.452
Human	ERNIE Bot	0.716*	0.09	0	0.47	0.961
	Chatgpt	0.648*	0.09	0	0.402	0.893
	DeepL	0.697*	0.09	0	0.452	0.942

Note: Based on estimated marginal means:
*. The mean difference is significant at the 0.05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Table 6 One-way ANOVA

Dependent Variable: Score

	Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Contrast	5.698	3	1.899	29.411	0	0.595
Error	3.875	60	0.065			

Note: F tests the effect of Type. This test is based on the linear independence of the estimated marginal means for comparison.

From Tables 5 and 6, it can be seen that $F(3, 60) = 29.411$, $P<0.05$, partial eta-squared = 0.595; there is a significant difference among the three types of machine translation (ERNIE Bot, Chatgpt, Deepl) and human translation. What's more, there exists a difference among the three machine translations, but there is no significant difference among the three machine translations.

4 Results and Discussion

In this section, by analyzing the scores of human translation and machine translation, the author will conduct specific case studies based on the classification of English passive sentences. In the process of comparative analysis, the author will objectively compare the data to avoid subjective factors influencing the analysis results.

4.1 Agentful Passives

The structure of agentful passive sentences is NP + BE + V-EN + BY + NP, which consists of both agent and patient and belongs to prototypical passives. The scores of machine translation and human translation in agentful passives are shown in the Table 7 and Table 8.

Table 7 The Scores of Machine Translation and Human Translation (Agentful Passives)

	Faithfulness	Fluency	Total
ERNIE Bot 1	4.33	4.44	4.37
Chatgpt 1	3.89	4.33	4.07
Deepl 1	4.11	4.44	4.24
Human 1	4.78	4.67	4.74
ERNIE Bot 2	3.89	4.33	4.07
Chatgpt 2	3.78	4.44	4.04
Deepl 2	4	4.56	4.22
Human 2	4.67	4.67	4.67
ERNIE Bot 3	4.33	4.33	4.33
Chatgpt 3	4	4.11	4.04
Deepl 3	3.78	4	3.87
Human 3	4.78	4.67	4.74
ERNIE Bot 4	3.78	4	3.87
Chatgpt 4	4	4.11	4.04
Deepl 4	4.22	4	4.13
Human 4	4.56	4.44	4.51

Table 8 The Average Scores of Machine Translation and Human Translation (Agentful Passives)

	Average of Faithfulness	Average of Fluency
ERNIE Bot	4.08	4.28
Chatgpt	3.92	4.25
Deepl	4.03	4.25
Human	4.7	4.61

The average scores for faithfulness and fluency of three types of machine translation are significantly lower than the average scores for faithfulness and fluency of human translation, which is analyzed further by considering specific examples.

Example 1:

Source Text: If you aren't convinced by my arguments that ethics is connected to cooperation problems, the rest of the book is still valuable.

ERNIE Bot: 如果你并不认同我的论点，即伦理与合作问题有关，那么本书剩余部分仍然有价值。

Chatgpt: 如果你对我认为伦理学与合作问题有关的论点并不信服，这本书的其余部分仍然具有价值。

Deepl: 如果你不相信我关于伦理与合作问题相关联的论点，本书的其余部分仍然很有价值。

Human: 我认为伦理与合作问题有关，如果你不这么认为，这本书的其余部分仍然值得一读。

In the source text, the agent in the clause "my arguments that ethics is connected to cooperation problems" is a noun phrase with an appositive clause. Since English gives priority to hypotaxis, to make the sentence more cohesive, passive construction is used to avoid the subject in the clause being too long. On the contrary, parataxis is a distinctive feature of Chinese, and Chinese sentences are more flexible. In example 1, human translation and machine translation follow the order of the passive sentence. The tip is that if there are both agent and patient in the original passive construction, it is often translated into a Chinese active sentence with the original agent as the subject and the original patient as the object. This is one of the basic methods to translate an English passive sentence.

4.2 Agentless Passives

The structure of agentless passive sentences is NP+BE+V-EN, which consists only of the patient and belongs to prototypical passives. The scores of machine translation and human translation in agentless passives are shown in Table 9 and Table 10.

Table 9 The Scores of Machine Translation and Human Translation (Agentless Passives)

	Faithfulness	Fluency	Total
ERNIE Bot 5	4	4.11	4.04
Chatgpt 5	4.22	4.22	4.22
Deepl 5	4.22	4.33	4.26
Human 5	5	5	5.00
ERNIE Bot 6	4.33	4.22	4.29
Chatgpt 6	4	4.11	4.04
Deepl 6	3.44	3.89	3.62
Human 6	5	5	5.00
ERNIE Bot 7	3.78	4.11	3.91
Chatgpt 7	4.33	4.44	4.37
Deepl 7	4.11	4.33	4.20
Human 7	4.67	4.78	4.71
ERNIE Bot 8	3.78	4	3.87
Chatgpt 8	4.33	4.56	4.42
Deepl 8	4	4.22	4.09
Human 8	5	4.89	4.96

Table 10 The Average Scores of Machine Translation and Human Translation (Agentless Passives)

	Average of Faithfulness	Average of Fluency
ERNIE Bot	3.97	4.11
Chatgpt	4.22	4.33
Deepl	3.94	4.19
Human	4.92	4.91

The average scores for faithfulness and fluency of three types of machine translation are significantly lower than the average scores for faithfulness and fluency of human translation, which is analyzed further by considering specific examples.

Example 2:

Source Text: But her letters were always long expected, and always very short.

ERNIE Bot: 但是她的信总是让人盼望很久，也总是很短。

Chatgpt: 但她的信总是被期盼很久，却总是非常简短。

Deepl: 但她的信总是期待已久，而且总是很短。

Human: （前文提到的人）总是久久盼着她的来信，而她的信却总是只有寥寥几句。

In the human translation version, the subject implied in the passive sentence is supplemented as the subject, while the other three machine translations are still translated according to the order of the passive sentence. This is where the difference lies between human translation and machine translation. The method of making the omitted agent the subject is usually applied to agentless passives. Since in English, the agent is usually omitted but explicit in meaning, it is usually figured out to be the subject in the Chinese translation. From this, it can be seen that when translating agentless passive sentences, machines still need to improve their ability to recognize the agent in the passive sentence by considering the context. In words, machine translation can accurately translate simple passive sentences, but there is still a certain gap between it and human translation in terms of context recognition and linguistic readability.

4.3 Stative Sentence

The structure of stative passive sentences is NP + BE (copular) + V-EN (adj.), which belongs to non-prototypical passives. The scores of machine translation and human translation in stative passives are shown in Table 11 and Table 12.

Table 11 The Scores of Machine Translation and Human Translation (Stative Sentence)

	Faithfulness	Fluency	Total
ERNIE Bot 9	4.44	4.56	4.49
Chatgpt 9	4.22	4	4.13
Deepl 9	4.11	4.33	4.20
Human 9	5	5	5.00
ERNIE Bot 10	3.78	3.89	3.82
Chatgpt 10	4	4.11	4.04
Deepl 10	3.89	4	3.93
Human 10	4.89	4.78	4.85
ERNIE Bot 11	4	4.11	4.04
Chatgpt 11	3.89	4	3.93

Continued

	Faithfulness	Fluency	Total
Deepl 11	3.56	3.44	3.51
Human 11	4.56	4.56	4.56
ERNIE Bot 12	3.33	3.78	3.51
Chatgpt 12	3.78	4	3.87
Deepl 12	3.44	3.67	3.53
Human 12	4.67	4.44	4.58

Table 12 The Average Scores of Machine Translation and Human Translation (Stative Sentence)

	Average of Faithfulness	Average of Fluency
ERNIE Bot	3.89	4.09
Chatgpt	3.97	4.03
Deepl	3.75	3.86
Human	4.78	4.69

The average scores for faithfulness and fluency of three types of machine translation are significantly lower than the average scores for faithfulness and fluency of human translation, which is analyzed further by considering specific examples.

Example 3:

Source Text: He was not willing to have the attempt known till its success could be known likewise.

ERNIE Bot: 直到事情成功，他才愿意让人知道他曾有过这次尝试。

Chatgpt: 他不愿意在成功之前就让人知道这个尝试。

Deepl: 在成功之前，他不愿意让人知道他的尝试。

Human: 先行其言，而后从之。

As a matter of fact, compared with"被", a great deal more English passives are translated into"给""叫""让". In example 3, three machines have all processed the passive voice and translated it as"让". But the linguistic readability of the three versions of machine translation is inferior to that of human translation. The reason is that human translators have subjective initiative; they can choose the wording and expressions that best match the target language culture based on their knowledge and experience, which is precisely the lack of machine translation.

4.4 Medio Passives

The structure of medio passive sentences is NP (person) + Copular Verb + V-EN + (Prep + NP) or NP + GET + V (get passives), which belongs to non-prototypical passives. The scores of machine translation and human translation in medio passives are shown in Table 13 and Table 14.

Table 13 The Scores of Machine Translation and Human Translation (Medio Passives)

	Faithfulness	Fluency	Total
ERNIE Bot 13	4.44	4.56	4.49
Chatgpt 13	4.33	4.44	4.37
Deepl 13	4.22	4.44	4.31

Continued

	Faithfulness	Fluency	Total
Human 13	4.89	5	4.93
ERNIE Bot 14	3.89	3.78	3.85
Chatgpt 14	4	4.11	4.04
Deepl 14	4.33	4.44	4.37
Human 14	4.78	4.89	4.82
ERNIE Bot 15	3.78	4	3.87
Chatgpt 15	4.22	4.44	4.31
Deepl 15	4.67	4.78	4.71
Human 15	5	5	5.00
ERNIE Bot 16	4.56	4.33	4.47
Chatgpt 16	4.45	4.44	4.45
Deepl 16	4.44	4.33	4.40
Human 16	4.67	4.67	4.67

Table 14 The Average Scores of Machine Translation and Human Translation (Medio Passives)

	Average of Faithfulness	Average of Fluency
ERNIE Bot	4.17	4.16
Chatgpt	4.25	4.36
Deepl	4.41	4.5
Human	4.84	4.89

The average scores for faithfulness and fluency of three types of machine translation are relatively lower than the average scores for faithfulness and fluency of human translation, among which Deepl performs best in machine translation and receives high recognition. The sentence with the lowest average score is analyzed below.

Example 4:

Source Text: Let’s say getting hit by a train will lead to a 99 percent chance of losing everything that’s important.

ERNIE Bot: 假设被火车撞击会导致 99% 的几率失去所有重要的东西。

Chatgpt: 假设被火车撞击会导致失去一切重要事物的可能性达到 99%。

Deepl: 比方说，如果被火车撞上，99% 的几率会失去所有重要的东西。

Human: 比方说，被火车撞了，那么人失去一切重要东西的几率将达到 99%。

In example 4, “getting hit by” implicates negative emotion in the context. Chinese passive voice has been called “inflective voice” and Chinese syntactic passive construction with marks like“ 被 ”“ 给 ”“ 挨 ”“ 遭 ”“ 让 ”“ 受 ”often can express such kinds of emotions. Here, all three machine translations do a good job. The reason why these three machine-translated sentences get low scores is that the three machine translations fail to make the omitted agent the subject, which leads to the sentence not conforming to the grammatical norms.

4.5 Reasons and Suggestions

In the translation of agentful passive sentences, machine translations have no major problems in terms of faithfulness and fluency, as it is a well-developed area. This is because both the agent and the patient are present in agentful passive sentences, so there is no need for machines to do the work of recognition and judgment. However, when it comes to translating agentless passive sentences, stative sentences, and medio-passives, there is a significant gap between machine and human translation. The main problems are that the translation methods used by machine translators are not flexible, and passive voice will be mechanically translated into the Chinese character " 被 " by machine translators. The reasons may include failing to recognize the agent or patient of the passive sentences, even confusing the agent and patient, and not having enough storage of Chinese characters to express passive voice, being unqualified to identify and select Chinese characters that best fit Chinese expression habits, and so on.

Therefore, it is recommended to strengthen the code strength of the logic of the agent and the patient, expand the storage capacity of machine translation in the area of Chinese characters that can convey passive voice, like" 被 "" 给 "" 挨 "" 遭 "" 让 "" 受 ", and complement the corresponding example sentences. What's more, when editing more effective instructions that can be recognized and executed by the machine to assist in deep learning and classification, the focus should be on providing clear guidelines for data processing, feature extraction, and model training. By defining specific steps and parameters in the instruction set, the machine can better understand how to analyze and categorize information, leading to improved decision-making and word selection accuracy. This approach enhances the machine's adaptability in selecting appropriate language and improves its overall performance in various tasks.

Machine translators are generally able to accurately translate simple passive sentences, but there is still a gap with human translations in terms of context recognition and language readability. Translation is not just a simple conversion between two languages, but is also influenced by many factors such as the language itself, the translator, and the reader.

On the one hand, from the perspective of the form of language itself (including word categories, syntactic structures, etc.), Nida (2012) believes that translation can have two orientations: "formal equivalence" and "dynamic equivalence". Equivalence in form and content is called "formal equivalence", while equivalence in cultural style is referred to as "dynamic equivalence". However, as no two languages can achieve complete equivalence in organization and meaning, absolutely accurate translation does not exist. Currently, machine translation is basically at the stage of word-for-word translation. For longer English passive sentences, machine translation often produces long and awkward translations due to literal word correspondences. In contrast, human translators will appropriately transform the voice based on Chinese thinking and expression habits, combined with the surrounding context, to provide a naturally fluent translation of passive sentences.

In comparison to the richness and diversity of natural language, computer rules are relatively limited and cannot cover all semantic phenomena. Therefore, for the translation of complex syntactic structures in specific domains, linguists and computer experts need to collaborate and make joint efforts to improve translation quality by simulating artificial neural network technology and machine deep learning, editing more effective instructions that can be recognized and executed by machines.

On the other hand, from the perspective of translators, culture, social environment, and other relevant factors, language is the carrier of culture, so human translators can choose the most culturally appropriate wording and expressions based on knowledge and experience. However, machine translation can perform word-for-word conversion and convey the most basic literal information, but there is still much room for improvement in conveying pragmatic information and cultural connotations in texts. Semantic equivalence is just a basic requirement of translation;

translations should also strive to be appropriate and natural, providing target language readers with the same feelings as the source language. Therefore, it is necessary to develop more sophisticated algorithms for context recognition and language modeling. Additionally, it is also important to train machine translation models with larger and more diverse datasets to increase their generalization ability and adaptability to different translation scenarios.

6 Conclusion

This paper takes English passive sentences as a concrete research area and compares the translations of human translators and three machine translators, including ERNIE Bot, ChatGPT, and Deepl. Through quantitative research, in translating English passive sentences, there is a significant difference between the three types of machine translation and human translation. What's more, there exists a difference among the three machine translations, but there is no significant difference among them. Based on this, qualitative analysis is further elaborated, including progress and shortcomings that machine translation possesses, the reasons for these shortcomings, and suggestions for improvement.

Due to various reasons, this paper has several limitations. Firstly, in terms of research corpus, the author did not focus the selection of English passive sentences on a specific field, such as technical English, business English, literary works, etc. And the number of selected English passive sentences was not sufficient, thus may lack representativeness of the results of this paper. Secondly, during the process of selecting the materials, the author consulted numerous corpora and literature, and made selections based on the classification criteria for passive sentences. However, the author's choices and matching during the classification process may have been influenced by her own level of knowledge, potentially leading to subjective biases in the final selection of materials. Thirdly, this paper involved a questionnaire survey, with a total of 27 valid responses collected. If there were a greater number of questionnaires, the resulting analysis might have been more ideal. The author will continually improve upon these issues in future research.

References

[1] Bai, X. J., & Zhan, W. D. (2003). Constraints on Chinese" 被 "sentences and processing of English passive sentences in machine translation. In *New expansions in the study of Chinese passive expressions: Proceedings of the International Symposium on Chinese Passive Expressions*. Wuhan, China: Huazhong Normal University Press.

[2] Gao, L. L., & Zhao, W. (2020). A review of machine translation research. *China Foreign Languages,* (6), 97-103.

[3] Halliday, M. A. K. (2008). A*n introduction to functional grammar (2nd ed.).* Beijing, China: Foreign Language Teaching and Research Press.

[4] Kempson, R. M. (1981). *Semantic theory. Cambridge*, United Kingdom: Cambridge University Press.

[5] Li, M., & Zhu, X. M. (2013). A new exploration of English-Chinese machine translation with post-editing automation. *Chinese Translators Journal, 34*(4), 83-87.

[6] Nida, E. (2012). *Principles of correspondence. In Selected readings on western translation studies*. Wuhan, China: Wuhan University Press.

[7] Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London, United Kingdom: Longman.

[8] Wang, G. Y. (1984). Generation of Chinese sentence patterns for English "be+Ved" passive sentences in automatic English-Chinese translation. *Studies in Language,* (2), 41-56.

[9] Wang, Z. J. (2004). *A cognitive contrastive study of English and Chinese passive sentences*. Shanghai, China: Shanghai Foreign Language

Education Press.

[10] Wang, Z. J., & Wei, H. (2002). On the derivation of English and Chinese passive sentence patterns. *Journal of Foreign Languages*, (3), 44-55.

Appendix

1. Cleaning and user maintenance shall not be made by children without supervision.

2. Whichever was the case, her opinion of him must be materially affected by the difference.

3. If you aren't convinced by my arguments that ethics is connected to cooperation problems, the rest of the book is still valuable.

4. The power of doing anything with quickness is always much prized by the possessor.

5. The challenge for a machine learning approach to designing robot ethics is that choices must be made about what kind of information is used as positive or negative feedback.

6. But her letters were always long expected, and always very short.

7. It was subsequently revealed that the Clipper Chip would enable law-enforcement officials to have access to data in the computer.

8. Smaller assemblages of similar pieces were also recovered from underlying layers that have not yielded any human bones.

9. He was not willing to have the attempt known till its success could be known likewise.

10. If the dispositions of the parties are ever so well known to each other or ever so similar beforehand, it does not advance their felicity in the least.

11. I am always glad to get a young person to be well placed out.

12. The rest of the evening was spent in conjuring how soon he would return Mr. Bingley's visit, and determining when they should ask him to dinner.

13. Although Americans today are likely to think that Alger's stories are too good to be true, they continue to be inspired by the ideas of earning wealth and success as an entrepreneur who makes it on his own.

14. Let's say getting hit by a train will lead to a 99 percent chance of losing everything that's important.

15. Sadly, the film is let down by an excessively simple plot.

16. Does that mean God is so great that he can't be bothered with caring for you and me?