Evaluation of Machine Translation in English-Chinese Automatic Subtitling of TED Talks

Guangyuan YAO[1]

[1]School of English Studies, Shanghai International Studies University, E-mail:guangyuanyao@shisu.edu.cn

Abstract

With technology becoming an essential competitive factor for subtitle translation services, the quality of machine subtitle translation in the current era deserves further examination. This paper examines the Chinese translation of 20 Ted Talks subtitled by an automatic subtitling software under the frame of the FAR model. This paper examines if machine subtitle translation has semantic, contextual, grammatical or spelling errors, idiomatic errors, syncopation and inappropriate synchronization, as well as wrong reading speed and paragraph length. It also emphasizes the need to broaden new perspectives on the quality of machine subtitle translation.

Key Words: machine translation, automatic subtitling, English-Chinese translation, translation technology, TED Talks

1. Introduction

In the era when data-driven statistical machine translation is dominant, scholars explore the application of machine translation in subtitling. With the development of speech-to-text technology and neural network machine translation, video subtitle synchronization translation technology has developed rapidly, and a number of scholars propose new methods to improve its accuracy. In China, the study of audiovisual translation continues to evolve and has become an important area of translation studies, receiving more attention. At the same time, China's domestic audiovisual market is booming and growing in influence. In the age of AI, as international exchange unfolds widely and intensively, speed is increasingly pursued, and to achieve this, the strong involvement of electronic technology, the emergence of language localization projects, the use of translation memories, the popularity of machine translation, and even the application of AI technologies with greater efficacy have all improved the efficiency of human communication to varying degrees. At the same time, the broad vision of introducing localization, corpus, project management, programming and other elements that were not previously part of the translation field is popular.

In the field of subtitle translation, time is of great importance and AI subtitle translation tools can contribute a lot to solving the time problem. In the age of globalized and intelligent information, one of the main challenges facing translation and cultural communication is the fact that the huge number of subtitle translation tasks exceeds the capacity of a limited number of translators, i.e. how to quickly translate the urgently needed film and television productions in a limited time frame. Some scholars point out that "there is an underlying criterion in the field of subtitle translation, that is, whoever can get the work on screen in the shortest time will win more attention and market" (Wang & Xi, 2014: 34).

With a short time frame, it is inevitable that problems such as inconsistent translations of front and back terms, haphazardly conducted or even ignored review and proofreading sessions, and frequent errors occur. In addition to time problems, fansubbers also have problems with the division of labor and collaboration between members, mainly due to differences in a professional level and time and place, which makes it difficult for members to coordinate their work effectively and for multiple people to work together, resulting in inconsistencies in the overall language style of the film or television production. Machine subtitle translation can complement human translation in terms of unified terminology, improved translation efficiency and clear division of labor, and has enormous advantages. However, the quality of machine-translated subtitles has not received enough attention in academics. "Little has been asked about new technological developments such as computer-aided translation software, terminology technology, and speech recognition technology in subtitle translation, and little research has been done on the speed efficiency and quality assurance of subtitle translation." (Wang & Xi, 2014: 34), so the quality of machine subtitle translation in the new era is worth exploring.

Besides, the quantitative methods are rarely adopted in the Chinese context in audiovisual translation field. Cao points out that "there are still many shortcomings in domestic Internet subtitle translation research, the most obvious of which is that a large proportion of articles still follow traditional translation theories to evaluate the translation quality of subtitles." (2017: 28) Throughout the local Chinese studies, most of them adopt the traditional translation theory rather than the assessment model specific to subtitle translation to explore translation quality. Huang and Zhang point out that domestic audiovisual translation research should "maintain the usual case studies and theoretical interpretations while exploring and using advanced empirical research tools."(2019: 85).

In 2017, Jan Pedersen, a Swedish scholar, proposed a model for assessing the quality of inter-lingual subtitling, the FAR model, filling a gap in assessing of inter-lingual subtitling quality. He argues that before the FAR model was proposed, subtitle quality was mainly assessed by subtitle norms and industry judgments, while the FAR model takes into account all factors and uses quantitative methods to calculate it. (2017: 214) In this model, F refers to Functional Equivalence and A refers to Acceptability while R refers to Readability. Also, Peterson (2017) uses the FAR model to study16 versions of Swedish subtitles for 10 English-language films to compare and analyses the differences between subtitling teams and official standard subtitles, which proved that the model could be adapted accordingly to the subtitling standards of different language regions.

Pedersen (2017) categorizes translation errors into three dimensions in the FAR model: functional equivalence, acceptability and readability. Functional equivalence is to convey both what is said and what is meant. Acceptability is used to assess the extent to which the translation conforms to the linguistic norms of the target language. Readability relates to the technical specifications of the translation of subtitles. The author sets the smallest unit of quality assessment at the subtitle itself (one to two lines), which not only ensures the integrity and relative independence of the meaning group, but also improves the operability of the assessment model. The formula for calculating the overall subtitle quality score is N-F-A-R/N*100, where N is the number of units, i.e. the number of subtitles, and F, A and R are the points deducted for the three main types of error.

A number of subtitle translation tools have been developed by domestic businesses, such as NetEaseSight, RenRen Translation Vision, SubtitleTone, etc. NetEaseSight was launched in May 2017, which is based on advanced Neural Network technology, combined with data mining, algorithmic recommendation and other capabilities to aggregate overseas content in one click, quickly translate the full text through machine translation and automatically generate subtitle files. On 28 May, 2018, RenRen Translation World was created and developed by the original RenRen Video founding team, combining multi-person collaboration management with AI video translation technology provided by NetEaseSight, which has 15 years of experience in video translation and has rich video translation experience and corpus resources. In addition to automatic subtitle generation, the software also provides a crowdsourcing translation platform, a translation collaboration platform, official translation services and video translation tools. Therefore, this paper selects the subtitles automatically generated by the NetEaseSight platform as the object of research and conducts a quality assessment.

In this paper, the author will select the subtitles automatically generated by the NetEaseSight platform as the object of research and conducts a quality assessment. The author uses the top 20 TED talks uploaded on the NetEaseSight platform to generate machine translated subtitles. The author examines the accuracy of speech recognition and cut scores, applying the FAR model to evaluate the quality of machine-translated Chinese subtitles.

Table 1. Top 20 TED Talks

Sequence Number	The Name of the Ted Talk	The length of the Talk	Key Words
1	Do Schools Kill Activity	19:08	Creativity; Culture; Dance
2	This Is What Happens When You Reply to Spam Email	09:52	Comedy; Curiosity; Communication
3	Your Body Language May Shape Who You Are	21:00	Body Language; Brain; Business
4	How Great Leaders Inspire Action	17:44	TEDx; Business; Entrepreneur
5	Inside the Mind of a Master Procrastinator	13:55	Brain; Comedy; Decision-Making
6	The Power of Vulnerability	20:04	TEDx; Communication; Culture
7	How to Speak So That People Want to Listen	09:41	Culture; Sound; Communication
8	My Philosophy for A Happy Life	12:45	Happiness; Personal; Growth; Self
9	The Next Outbreak We're Not Ready	08:20	Disease; Ebola; Global Issues

10	What Makes A Good Life Lessons from The Longest Study on Happiness	12:37	TEDx; Aging; Data
11	Looks Aren't Everything. Believe me, I'm A Model.	09:22	TEDx; Beauty; Culture
12	Why People Believe They Can't Draw and How to Prove They Can	15:03	Art; Design; Potential
13	10 Things You Didn't Know About Orgasm	16:24	Culture; History; Humor
14	The Orchestra in My Mouth	11:42	TEDx; Creativity; Music
15	How to Spot A Liar	18:30	Culture; Psychology; Science
16	The Art of Misdirection	08:30	Crime; Entertainment
17	The Power of Introverts	18:44	Business; Culture; Psychology
18	How I Held My Breath for 17 Minutes	20:16	Biology; Magic; Medicine
19	The Danger of a Single Story	18:29	Africa; Culture; Identity
20	My Stroke of Insight	18:38	Biology; Brain; Consciousness

2. Literature Review

2.1 Previous Research on Machine-Translated Subtitles

Machine subtitling has been around since 2000 when Popowich (2000) built a fully automated, large-scale multilingual machine translation system, ALTo, using interpretation-based learning techniques. The system is evaluated in three ways: correct; acceptable though unsatisfactory; unacceptable. The system was evaluated manually in English-Spanish pairs with an accuracy of 70-80%. Armstrong (2006) conducts a real-user evaluation of an example-based machine translation (EBMT) of English subtitles into German and Japanese and presents the advantages of machine translation and identifies a number of issues that need to be addressed, including the need for more creativity when restructuring translations and the problem of sentence segmentation.

Volk (2008) investigates rule-based, instance-based and statistical-based machine subtitle translation respectively and concludes that the machine subtitle translation system is capable of producing passable subtitles. Sousa (2011) conducts experiments to evaluate a phrase-based SMT system for English Portuguese using the Moses toolkit and Google Translate, and scored the quality of the output using the Bilingual Evaluation System (BLEU), with the final results showing that post-editing of pre-translated subtitles by subtitlers would be less time-consuming. Aziz (2012) assesses the quality of caption generation using Google Translate built by Moses and two phrase-based statistical machine translation systems, calculated using the BLEU, Translation Editing Rate (TER) metric, presented a positive attitude toward machine translated captions.

Etchegoyhen (2014) conducts experiments based on approximately one million aligned parallel subtitles and 15 million monolingual subtitles in 14 bidirectional language pairs, using BLEU and TER scores to assign a positive rating to each machine-translated subtitle and a survey with user feedback to assess the quality of the machine-translated subtitles. Despite the positive comments on the quality of machine subtitle translations, not all of these studies have been well received. O Hagan (2003) compares English-Japanese online machine-translated subtitles (powered by the Amikai translation engine) for The Lord of the Rings with human-translated Japanese subtitles and showed that the actual match between a translation memory and translated text and the repetition of translated text was not satisfying. Diaz-Cintas and Remael (2014) conclude that it will be several years before machine-translated subtitles reach a sufficient level of linguistic quality to fully meet the needs of translation consumers.

Although foreign scholars have long been involved in research on machine subtitling translation, there is still a large gap in this area in China, and there are only a few articles in the domestic literature at present. Wang and Xi point out that "although the research on the direction of film and television subtitle translation in China has become more in-depth in the past ten years, most of the research results still mainly focus on the main problems in the translation process and the corresponding translation strategies, etc." (2014: 33). It can be seen that most studies are conducted from a traditional perspective while little research

has been conducted on some new topics such as computer-aided translation software, terminology technology, and speech recognition technology.

Wang and Xi (2014) analyze the problems faced by subtitle translation and focus on explaining that computer-aided translation technology can help unify terminology in subtitle translation by extracting high-frequency terms (names of people, places, objects, organizations, etc.) and making a terminology database that can be shared by translation team members, saving time and labor, and also ensuring the uniformity and consistency of the translation of the same terminology in the whole film and television work. The article illustrates the extremely important role played by computer-aided translation technology in improving the quality and efficiency of subtitle translation, and concludes by pointing out that computer-aided translation technology is of great significance in promoting cultural communication, enriching the content of subtitle translation teaching and expanding the scope of current translation studies.

Yu (2019) conducts an experiment on machine subtitle translation, based on two videos of lectures on design thinking and water resources, with English as the main language and English to Chinese subtitle translation. The thesis concludes that the AI (Artificial Intelligence) hyped by subtitle translation platform vendors cannot produce high quality subtitle translations on their own for the time being, while the human-computer interaction subtitle translation model of machine translation and human proofreading can greatly improve the speed of subtitle translation while ensuring the quality of subtitles. However, the paper still has shortcomings: the selection of corpus content is relatively single topic; only two test platforms are selected; only a typical case is selected by individuals, no acceptance study is conducted, which is too subjective.

A recent study by Xiao and Gao (2020) is more cutting-edge, using the "FAR model" constructed by Pedersen to evaluate the quality of subtitles automatically generated by "NetEaseSight" from English to Chinese subtitles and NetEase Open Classes using TED talks as the source material. The results show that there is still a large gap between the two. By combining the advantages of machine and human, the development of machine subtitle translation is more promising. It can be said that this paper is groundbreaking. However, the paper also has a slight flaw in that the sample is a 14 minute 8 second video of a TED talk, which is a somewhat small sample size. If the sample size can be expanded by this research path, it can provide more research space and development possibilities for Chinese audiovisual translation research.

Wang and Li (2020) argue that artificial intelligence has empowered translation technology, providing basic and global technical support for subtitle translation, and the importance and urgency of research on subtitle translation technology is becoming increasingly prominent. They use literature analysis to sort out the current status of research on subtitle translation technology, analyze its main problems, explore future development trends, and put forward targeted suggestions to dovetail with national cultural communication It also proposes recommendations to expand the research on subtitle translation technology, strengthen education on subtitle translation technology and promote the innovation mechanism of integration of government, industry, academia, research and application in the field of subtitle translation.

2.2 Previous Research on Quality Assessment of Subtitles

The seminal research on subtitling quality assessment can be traced back to Brondeel's (1994) study. He examines three levels of equivalence of subtitling, arguing that subtitling must be closely approximate to the spoken source language in relation to discourse. Karamitroglou (1998) suggests that the practice of television subtitling must maximize the enjoyment and understanding of the target film by making the inserted text as readable and accessible as possible. Gottlieb distinguishes between subtitling and other forms of translation in terms of four simultaneous symbolic channels (i.e. image, sound, dialog and subtitles) from a multimodal perspective and argues that examining the subtitler's language goal and camera movements first insteadof individual lexical pieces affords them linguistic freedom. As a result, it is vital to analyze the subtitle as a whole and strive to recover the semantics of the original work while analyzing the quality of subtitling. (2001: 19)

Diaz-Cintas (2001) focuses on the subtitling decision-making process, arguing that because subtitlers prefer to use other semiotic qualities of the film, the quality of an audiovisual translation must be assessed by subtitlers. (2001: 205) To measure the effectiveness of dialog translation, he summarizes five major elements of subtitling, which are discourse cohesion and coherence, subtitling to achieve the same effect as the original, reproduction of different linguistic systems, implementation of compensatory strategies, and the absence of defects or errors. Bittner (2011) suggests six factors that influence the translation process, including: textual form, culture, translator, source text, politics and client. He provides a detailed grammatical and contextual analysis of the subtitles of Chapter 14 of the Murder on the Orient Express film. The paper points out the importance of the quality of subtitles in film as an art form, as they are an integral part of the film.

Pedersen (2017) proposes three criteria for evaluating the quality of interlingual subtitles: whether the subtitles convey the speaker's meaning, whether the subtitles are correct and natural in the target language, and whether the subtitles can be read fluently and without interruption. He explores the quality indicators by which professional subtitle translators and viewers judge subtitling, exploring how influencing factors such as letter breaks, synchronization and display speed affect subtitle

quality, and ultimately concluding that professional subtitle translators and viewers have different attitudes to subtitle condensation and that different groups judge subtitle quality differently.

Wang and Li (2020) point out that the development of speech recognition technology has made the research on quantitative assessment of audiovisual translation a new hotspot in the current international translation research. The current situation of quantitative assessment of audiovisual translation in foreign countries is systematically reviewed, and it is found that audiovisual translation products are the main object of quantitative assessment research. The existing assessment models mostly adopt the error deduction system. Besides, pluralism and communicative nature are the characteristics of quantitative assessment research of audiovisual translation. For the international research trends and the current situation of domestic research, the assessment theory construction and quantitative model application of audiovisual translation may become a new direction for future research.

At the same time, in addition to the above-mentioned subtitle translation assessment criteria, quality assessment of machine subtitle translations might take the form of various indications. Machine subtitle translation quality can be assessed either automatically or by human assessors. Automatic evaluation was created to lower the cost and time of assessment because manual evaluation is time demanding. BLEU, National Institute of Standards and Technology (NIST), Translation Edit Rate (TER), and Word Error Rate are the five basic automatic evaluation metrics used to evaluate machine translated subtitles (WER). The number for BLEU (Papineni, 2002) is 0 to 1, with higher scores reflecting better translations. The bottom bound is 0 in NIST (Doddington, 2002), but there is no theoretical limit, with higher scores reflecting better translations. The WER (Och, 2003) scale ranges from 0 to 1, with lower values indicating better translations. With better translations, TER (Snover, 2005) has a lower score. The lower the word error rate (WER) indication, the better the translation.

Audiovisual translation studies are expanding in today's world, and it has become an essential subject of translation studies that is garnering greater academic attention. Simultaneously, China's domestic audiovisual market is booming, and its influence is growing, thus the quality of subtitle translation in the Chinese context is worth investigating. However, for domestic studies in China, most of them use traditional translation theories to investigate the quality of translation rather than assessment models specific to subtitle translation, and there are only a few quantitative studies that use special subtitle assessment models like the FAR model. This paper will conduct a quantitative analysis of the machine subtitle translation of TED Talks based on the FAR model from the three criteria introduced by the model: functional equivalence; acceptability; readability.

3. Results and Analysis of automatic subtitling from the FAR model

3.1 Results

Table 2. Score Calculation of Automatic Subtitling of Top 20 TED Talks

Sequence Number	The Name of the Ted Talk	Functional Equivalence Penalty Score	Acceptability Errors Penalty Score	Readability Errors Penalty Score	Number of Subtitles
1	Do Schools Kill Activity	-74	-2	-43.75	164
2	This Is What Happens When You Reply to Spam Email	-49	0	-15.15	82
3	Your Body Language May Shape Who You Are	-16	-2	-44.5	193
4	How Great Leaders Inspire Action	-13.5	-0.5	-32.5	170
5	Inside the Mind of a Master Procrastinator	-9	-0.75	-21.75	128
6	The Power of Vulnerability	-25.5	-0.5	-24.5	187
7	How to Speak So That People Want to Listen	-12.5	-1	-16.25	94
8	My Philosophy for A Happy Life	-5.5	-0.75	-7.75	111
9	The Next Outbreak We're Not Ready	-4.5	-0.25	-6	80

10	What Makes A Good Life Lessons from The Longest Study on Happiness	-4	-0.25	-8.5	112
11	Looks Aren't Everything. Believe me, I'm A Model.	-10.5	-0.25	-17	88
12	Why People Believe They Can't Draw and How to Prove They Can	-14	-0.5	-18.75	136
13	10 Things You Didn't Know About Orgasm	-18	-0.5	-18.5	145
14	The Orchestra in My Mouth	-9.5	0	-1.5	114
15	How to Spot A Liar	-17.5	-1.25	-35.25	173
16	The Art of Misdirection	-16.5	-1.25	-21.5	73
17	The Power of Introverts	-13.5	-1.5	-41.25	163
18	How I Held My Breath for 17 Minutes	-17.5	0	-41.25	183
19	The Danger of a Single Story	-17	0	-17	174
20	My Stroke of Insight	-14.5	0	-16.25	165

Table 3. Score of Automatic Subtitling of Top 20 TED Talks

Sequence Number	The Name of the Ted Talk	The Score of Functional Equivalence	The Score of Acceptability	The Score of Readability	The Final Score
1	Do Schools Kill Activity	54.14	98.77	73.16	26.38
2	This Is What Happens When You Reply to Spam Email	39.51	100	81.17	20.68
3	Your Body Language May Shape Who You Are	91.71	98.96	76.94	67.72
4	How Great Leaders Inspire Action	92.06	99.71	80.88	72.65
5	Inside the Mind of a Master Procrastinator	92.97	99.41	83.01	75.39
6	The Power of Vulnerability	86.36	99.73	86.90	72.99
7	How to Speak So That People Want to Listen	86.70	98.94	82.71	68.35
8	My Philosophy for A Happy Life	95.05	99.32	93.02	87.39
9	The Next Outbreak We're Not Ready	94.38	99.69	92.5	86.56
10	What Makes A Good Life Lessons from The Longest Study on Happiness	96.43	99.78	92.41	88.62
11	Looks Aren't Everything. Believe me, I'm A Model.	88.07	99.72	80.68	68.47
12	Why People Believe They Can't Draw and How to Prove They Can	89.71	99.63	86.21	75.55
13	10 Things You Didn't Know About Orgasm	87.59	99.66	87.24	74.48
14	The Orchestra in My Mouth	91.67	100	98.68	90.35
15	How to Spot A Liar	89.88	99.28	79.62	68.79
16	The Art of Misdirection	77.40	97.95	70.55	46.23
17	The Power of Introverts	91.72	99.08	74.69	65.49

18	How I Held My Breath for 17 Minutes	90.44	100	88.11	78.55
19	The Danger of a Single Story	90.23	100	90.23	80.46
20	My Stroke of Insight	91.21	100	90.15	81.36

3.2 Analysis

3.2.1 Functional Equivalence

Functional equivalence in the FAR model refers to that a subtitle would convey both what is said and what is meant. Thus, the errors are analyzed from both the saying and the meaning. The penalty points for semantic equivalence are minor: 0.5, standard: 1, and serious: 2. Pedersen states that "A serious semantic equivalence error scores 2 penalty points and is defined as a subtitle that is so erroneous that it makes the viewers' understanding of the subtitle nil and would hamper the viewers' progress beyond that subtitle." (2017: 219) However, in the model, it does not detail about the minor and standard errors much. In this case, a standard error means an error that makes the sentence unintelligible but does not hamper the whole subtitle. A minor error refers to an error that does not convey the word-level functional equivalence.

Translation Example 1:

> Source: Do Schools Kill Activity
> ST (Source Text): They have become frightened of being wrong. And we run our companies like this. We stigmatize mistakes.
> MST(Machine Source Text): They have become frightened of being wrong, and we run our companies. This, by the way, we stigmatize mistakes.
> MTT (Machine Translated Text): 他们害怕犯错 而我们管理自己的公司 顺便说一下 我们对错误进行污名化
> MTT in pinyin: tā men hài pà fàn cuò ér wǒ men guǎn lǐ zì jǐ de gōng sī shùn biàn shuō yī xià wǒ men duì cuò wù jìn xíng wū míng huà
> BT (Back Translation): They're afraid to make mistakes and we run our own company. By the way, we stigmatize mistakes.

The speaker wants to put forward that in the company's operations they have such problems, but there is no link in the translation, which causes the audience to hardly understand this sentence. In the translation, it is indicated that every person has a company, which contradicts the fact that there is only one company in ST. The example of companies shows the stigmatized wrong understanding of making mistakes. This belongs to the serious mistakes for the reason that this hinders the understanding of the particular sentence and the whole context. This mistake may be caused by the translation engine.

Translation Example 2:

> Source: This Is What Happens When You Reply to Spam Email
> ST: Now, my hand was kind of hovering on the delete button, right? I was looking at my phone. I thought, I could just delete this. Or I could do what I think we've all always wanted to do.
> MST: Now my hand was kind of hovering on the delete button, right as you are, I was looking at my phone, I thought I could just delete this, or I could do what I think we've all.
> MTT: 我的手好像停在删除按钮上 就像你一样 我看着我的手机 我想我可以直接删除这个 或者我可以做我认为我们都做过的
> MTT in pinyin: wǒ de shǒu hǎo xiàng tíng zài shān chú àn niǔ shàng jiù xiàng nǐ yí yàng wǒ kàn zhe wǒ de shǒu jī wǒ xiǎng wǒ kě yǐ zhí jiē shān chú zhè gè huò zhě wǒ kě yǐ zuò wǒ rèn wéi wǒ men dōu zuò guò de
> BT: My hand seems to be resting on the delete button, just like yours, and I'm looking at my phone and I think I could just delete this, or I could do what I think we both did.

This subtitle is translated wrongly for the reason that in the following sentence, the speaker is trying to telling the things she always wants to do. The things haven't been done but, in the translation, the things are expressed to have been done in the past. This may confuse the audience's feeling of time and the chronological order of the speech. This distorts the semantic meaning of the sentence, but does not affect the audience's understanding of the above, which can be called a standard error. This error is caused by wrong segmentation of MST.

Translation Example 3:

> Source: The Power of Introverts
> ST: When I was nine years old, I went off to summer camp for the first time. And my mother packed me a suitcase full of books.

MST: When I was nine years old, I went off to summer camp for the first time, and my mother packed me his suitcase full of books.

MTT: 当我九岁的时候 我第一次去夏令营 我的母亲给我装了他的箱子装满了书

MTT in pinyin: dāng wǒ jiǔ suì de shí hòu wǒ dì yī cì qù xià lìng yíng wǒ de mǔ qīn gěi wǒ zhuāng le tā de xiāng zi zhuāng mǎn le shū

BT: When I was nine years old, I went to my first summer camp and my mother packed me his suitcase full of books.

The original sentence means that when the speaker went off to the summer camp for the first time, his mother helped him pack his suitcase. However, in the translation, the ownership of the suitcase is incorrectly translated, which confuses the reader about the logic of the story. This minor error is caused by the error of the speech recognition.

3.2.2 Acceptability

Pedersen (2017) argues that acceptability refers to how closely the target text adheres to the target language's norms. This category contains faults that make the subtitles sound foreign or otherwise odd. These flaws can sabotage the illusion contract by drawing attention to the subtitles. Pedersen points out that "These errors are of three kinds: 1) grammar errors 2) spelling errors, 3) errors of idiomaticity." (2017: 220)

Translation Example 4:

Source: Your Body Language May Shape Who You Are

ST: In a few minutes, and I'm hoping that if you learned to tweak this a little bit, it could significantly change the way your life unfolds.

MST: In a few minutes, and I'm hoping that if you learned to tweak this a little bit, it could significantly change the way your life unfolds.

MTT: 在几分钟内 我希望如果你学会稍微调整一下 它可以极大地改变你的生活展开的方式

MTT in pinyin: zài jǐ fēn zhōng nèi wǒ xī wàng rú guǒ nǐ xué huì shāo wēi tiáo zhěng yī xià tā kě yǐ jí dà dì gǎi biàn nǐ de shēng huó zhǎn kāi de fāng shì

BT: In a few minutes I hope that if you learn to tweak it a little it can dramatically change the way your life unfolds

In Chinese "life" and "unfold" do not appear together. "Unfold" means "to become open" and in this context, it means life goes on. This is an error of idiomaticity.

Translation Example 5:

Source: How Great Leaders Inspire Action

ST: And I love asking businesses, what's your conversion on new business And they love to tell you, oh, it's about 10% proudly, well, you can trip over 10% of the customers.

MST: And I love asking businesses, what's your conversion on new business And they love to tell you, oh, it's about 10% proudly, well, you can trip over 10% of the customers.

MTT: 我喜欢问企业 你在新业务上的转化率是多少 他们喜欢自豪地告诉你 哦 大约是 10% 好吧 你可以绊倒超过 10%的客户

MTT in pinyin: wǒ xǐ huān wèn qǐ yè nǐ zài xīn yè wù shàng de zhuǎn huà lù shì duō shǎo tā men xǐ huān zì háo dì gào sù nǐ ó dà yuē shì 10 hǎo ba nǐ kě yǐ bàn dào chāo guò 10 de kè hù

BT: I like to ask businesses what your conversion rate is on new business, and they like to proudly tell you, oh, it's about 10%, well, you can trip up more than 10% of your customers.

To collocate with customers, the word "trip" cannot be directly translated into Chinese. It means you can outperform some customers in this case. The use of the word "trip" in Chinese is an idiomatic error.

3.2.3 Readability

Readability errors mean that viewers cannot read the subtitle effortlessly. In the FAR model, there are three parameters including segmentation and spotting; punctuation and graphics; reading speed and line length. In the results, errors of segmentation and spotting and reading and line length are found.

Translation Example 6:

Source: Your Body Language May Shape Who You Are

ST: So how many of you are sort of making yourself smaller Maybe you're hunching, crossing your legs,   maybe wrapping your ankles Sometimes we hold onto our arms like this.

MST: So how many of you are sort of making yourself smaller Maybe you're hunching, crossing your legs, maybe   wrapping your ankles Sometimes we hold onto our arms like this.

MTT: 你们中有多少人在把自己变小也许你们在驼背 交叉双腿 也许裹着脚踝有时我们像这样抱着胳膊

MTT in pinyin: nǐ men zhōng yǒu duō shǎo rén zài bǎ zì jǐ biàn xiǎo yě xǔ nǐ men zài tuó bèi jiāo chā shuāng tuǐ yě xǔ guǒ zhe jiǎo huái yǒu shí wǒ men xiàng zhè yàng bào zhe gē bó

BT: How many of you are making yourselves smaller, maybe you're hunching over, crossing your legs, maybe wrapping your ankles, sometimes we're holding our arms like this.

There are no spaces in the sentences in this translation and the segmentation is hard for the audience to follow.

In the FAR model, it only discusses that having too long lines will be difficult for the viewers to follow. However, a specific guideline for Chinese-English pair is not drawn up. According to Xiao (2020), translation over 30 characters in the target language can be called a standard error and translation over 20 characters but below 30 in the target language can be called a minor error. In this case, translation over 40 characters is defined as a serious error.

Translation Example 7:

Source: This Is What Happens When You Reply to Spam Email

ST: Now my hand was kind of hovering on the delete button, right as you are, I was looking at my phone, I thought I could just delete this, or I could do what I think we've all.

MST: Now my hand was kind of hovering on the delete button, right as you are, I was looking at my phone, I thought I could just delete this, or I could do what I think we've all.

MTT: 我的手好像停在删除按钮上 就像你一样 我看着我的手机 我想我可以直接删除这个 或者我可以做我认为我们都做过的 (50 characters)

MTT in pinyin: wǒ de shǒu hǎo xiàng tíng zài shān chú àn niǔ shàng jiù xiàng nǐ yí yàng wǒ kàn zhe wǒ de shǒu jī wǒ xiǎng wǒ kě yǐ zhí jiē shān chú zhè gè huò zhě wǒ kě yǐ zuò wǒ rèn wéi wǒ men dōu zuò guò de

BT: My hand seems to be resting on the delete button, just like yours, and I'm looking at my phone and I think I could just delete this, or I could do what I think we both did.

Translation Example 8:

Source: Looks Aren't Everything. Believe me, I'm A Model.

ST: I am on this stage, because I am a pretty white woman in my industry, we call that a sexy girl.

MST: I am on this stage, because I am a pretty white woman in my industry, we call that a sexy girl.

MTT: 我站在这个舞台上 因为在我的行业里 我是一个漂亮的白人女性 我们称之为性感女孩 (36 characters)

MTT in pinyin: wǒ zhàn zài zhè gè wǔ tái shàng yīn wèi zài wǒ de háng yè lǐ wǒ shì yí gè piào liàng de bái rén nǚ xìng wǒ men chēng zhī wéi xìng gǎn nǚ hái

BT: I'm standing on this stage because in my industry I'm a beautiful white woman who we call a sexy girl.

Translation Example 9:

Source: The Danger of a Single Story

ST: Is how impressionable and vulnerable we are in the face of a story, particularly as children.

MST: Is how impressionable and vulnerable we are in the face of a story, particularly as children,

MTT: 我们在面对一个故事时是多么的易受影响和脆弱 尤其是作为孩子 (28 characters)

MTT in pinyin: wǒ men zài miàn duì yí gè gù shì shí shì duō me de yì shòu yǐng xiǎng hé cuì ruò yóu qí shì zuò wéi hái zi

BT: How impressionable and vulnerable we can be when faced with a story, especially as children

4. Conclusion

The age of Artificial Intelligence has arrived with the advancement of information technology and computer hardware, and AI is employed in the form of Machine Translation (MT) to carry out translation duties. The usage of speech recognition translation technology has given rise to the creation of machine subtitling translation as machine translation continues to evolve. More and more translation is done by AI translation as the speed and quality of translation improves thanks to efficient machine translation facilitated by cloud-based systems. Machine-translated subtitles are generated by a large number of free and open-source tools in the industry. Although machine translation is quick and inexpensive, its accuracy is not always guaranteed.

Several suggestions can thus be made to improve the translation quality of machine-translated subtitles. The accuracy of machine translation engines should be improved. Low-frequency and unusual terms, for example, are deleted from the machine translation engine during the training phase to minimize the model's complexity and conserve storage space. The translation accuracy of some words should be improved. Some words sometimes aren't in dictionaries and are used to name things like persons, places, organizations, and trademarks, as well as time, dates, numbers, and new words. Furthermore, despite the fact that the translation is quite fluid, readability is still a huge issue. A correct structural examination of the source text is essential to fundamentally tackle the problem of order in translations. Semantic analysis is the foundation and

requirement for correctly analyzing a sentence's grammatical structure. However, there is a scarcity of substantial research achievements in this field, necessitating a greater emphasis on its application.

Finally, the accuracy of speech recognition engines and the function of segmentation should be improved. Although the technical maturity of voice recognition has substantially advanced, it is still not possible to attain one hundred percentage accuracy. Speech interaction is influenced by a variety of elements such as background noise and speech speed, and recognition rates vary widely from scene to scene; speech recognition lacks the ability to amend text through context; and insufficient semantics are the key bottlenecks. To solve this problem, the algorithm must be improved as well as a huge amount of reliable data for algorithm training in order for the algorithm to reach a particular level of maturity. However, it should be noted that although machine translated subtitles are still not of sufficient quality to go directly to market and require post-translation editing, they are accurate in a certain degree, with no grammatical mistakes or missing words. Furthermore, the automatically generated timetable is saving subtitlers' time and enhancing efficiency.

The future of subtitle translation services will be more promising if post-translation editing and machine translation of subtitles can be effectively combined. This study aims to provide some insights for the industry to improve the quality of machine subtitle translation, as well as to provide theoretical support for the industry to adopt machine subtitle translation, and also to broaden new perspectives for the quality of machine subtitle translation in China.

References

Armstrong, S., Caffrey, C., & Flanagan, M. (2006). Translating DVD subtitles from English-German and English-Japanese using example-based machine translation. *In MuTra 2006-Audiovisual translation scenarios: Conference proceedings.*

Aziz, W., de Sousa, S. C., & Specia, L. (2012). Cross-lingual sentence compression for subtitles. *In Proceedings of the 16th Annual conference of the European Association for Machine Translation*, 103-110.

Bittner, H. (2011). The quality of translation in subtitling. *Trans-kom*, 4(1), 76-87.

Brondeel, H. (1994). Teaching subtitling routines. *Meta: journal des traducteurs/Meta: Translators' Journal*, 39(1), 26-33.

Bywood, L., Etchegoyhen, T., Georgakopoulou, P., Fishel, M., Jiang, J., Loenhout, G. V., ... & Maucec, M. (2014, May). Machine translation for subtitling: A large-scale evaluation. *In LREC 2014, Ninth International Conference on Language Resources and Evaluation*, 46-53.

Cao, Y. X. (2017). Research Methods of subtitle Translation in the Internet Era: Current situation, Reflection and prospect. *Shanghai translation*, (5), 27-31. [In Chinese, 曹艺馨. (2017). 网络时代字幕翻译研究方法: 现状, 反思与展望. 上海翻译.]

Cintas, J. D. Striving for Quality in Subtitling: The Role of a Good Dialogue List, Multi (media) Translation Publishing House 2001. *Eds. Yves Gambier and Henrik Gottlieb*

Cintas, J. D., & Remael, A. (2014). Audiovisual translation: subtitling. *Routledge.*

Doddington, G. (2002). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. *In Proceedings of the second international conference on Human Language Technology Research*, 138-145.

Gottlieb, H. (2001). Texts, translation and subtitling-in theory, and in Denmark. *Translators and Translations*, 149-192.

Huang, Z. & Zhang, W. (2019). After Subtitle Translation: A review of Language and Culture Representation in Audiovisual Translation. *Oriental translation*, (06), 82-85. [in Chinese, 黄增& 张威. (2019). 字幕翻译之后——《视听翻译的语言与文化再现》评介. 东方翻译.]

Karamitroglou, F. (1998). A proposed set of subtitling standards in Europe. *Translation journal*, 2(2), 1-15.

O'Hagan, M. (2003). Can language technology respond to the subtitler's dilemma?-A preliminary study. *In Proceedings of Translating and the Computer 25.*

Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). Bleu: a method for automatic evaluation of machine translation. *In Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 311-318.

Pedersen, J. (2017). The FAR model: assessing quality in interlingual subtitling. *Journal of Specialised Translation*, (28), 210-229.

Popowich, F., McFetridge, P., Turcato, D., & Toole, J. (2000). Machine translation of closed captions. *Machine Translation*, 15(4), 311-341.

Snover, M., Dorr, B., Schwartz, R., Makhoul, J., Micciulla, L., & Weischedel, R. (2005). A study of translation error rate with targeted human annotation. *In Proceedings of the 7th Conference of the Association for Machine Translation in the Americas (AMTA 06)*, 223-231.

Volk, M., Nivre, J., Dahllöf, M., & Megyesi, B. (2008). The automatic translation of film subtitles: a machine translation success story?. *Studia Linguistica Upsaliensia*, 202-214.

Wang, H. S. & Li, Y. (2020). Research on subtitle translation technology: Current situation, problems and suggestions. *Audio-visual teaching of foreign languages*, (06), 80-85+6. [In Chinese, 王华树&李莹. (2020). 字幕翻译技术研究: 现状、问题及建议. 外语电化教学.]

Wang, H. S. & Xi, W. T. (2014). Research on subtitle Translation from the perspective of computer aided Translation technology. *English teacher*, 14(12), 32-38. [In Chinese, 王华树&席文涛. (2014). 计算机辅助翻译技术视角下的字幕翻译研究. 英语教师.]

Xiao, W. Q. &Gao, J. H. (2020). A study on Machine Translation Subtitle Quality Assessment: A case study of Chinese subtitle translated by NetEase. *Foreign languages and cultures*, 4(3), 95-105. [In Chinese, 肖维青&高佳晖. (2020). 机器翻译字幕质量评估研究——以 "网易见外" 英译中字幕为例. 外国语言与文化.]

Yu, F. Q. (2019). An experimental report on quality assessment of English-Chinese machine translation subtitles. *Southwest Jiaotong University*. [In Chinese, 余郁棋. (2019). 英汉机器翻译字幕质量评估实验报告. 西南交通大学.]