The Future of English-Chinese Translators and Interpreters in the Era of Artificial Intelligence

Approved: _____Richard L. Barrett___ Date: ___May 11, 2018_____
Paper Advisor

The Future of English-Chinese Translators and Interpreters in the Era of Artificial Intelligence

A Seminar Paper

Presented to

The Graduate Faculty

University of Wisconsin-Platteville

In Partial Fulfillment

Of the Requirement for the Degree

Master of Science in Education

English Education

By

CHEN OUYANG

2018

ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere gratitude to the people who have helped me throughout this seminar paper writing.

I sincerely appreciate Dr. Richard Garrett for his constructive instruction and earnest attitude in tutoring my process of this study. I would also thank Dr. Hu for her advise on my topic and materials she shared.

I wish to thank UWP faculties and friends who helped me in all respects.

Abstract

The Future of English-Chinese Translators and Interpreters in the Era of Artificial Intelligence

CHEN OUYANG

Under the Supervision of DR. GARRETT

Artificial Intelligence (AI), a widely discussed buzzword, approaches our life with acclamations as well as fears, showing its potential to significantly reshape the translation field. This paper reviews the notable breakthrough of new applications and experiments, followed by the analysis of human translators' threats, strength, weakness and opportunities. It then predicts the future human translators. At the end, it gives constructive suggestion to human translators in the coming era of AI.

TABLE OF CONTENTS

Chapter I Introduction

Artificial intelligence (AI) is greedily learning and reforming, gradually making the science fiction threats a reality. The technological singularity----when AI surpasses biological brains, is leading to an unimaginable revolution in human affairs. AI is spreading to thousands of domains, and it may eliminate many jobs like lawyers, bank clerks, sales representatives, etc. With Microsoft's recent announcement that their machine translation service can translate accurate and high-caliber sentences at human translators' level, it may feel as though the human translation industry is on the cusp of significant disruption. This paper will focus on the domain of English-Chinese translation.

Machine translation (MT), also known as Computer Aided Translation, is the use of software programs which have been specifically designed to translate both verbal and written texts from one language to another. Fiederer and O'Brien (2009) made an affirmative conclusion out of a set of experimental statistics, and they put forward that evidence shows when MT are used intelligently, it does not have to be synonymous with inferior quality translation. Compared with print translation, Zhang and Sun (2017) argued that Machine Translation should provide background information and dynamic images to reach the most comprehensive effects. MT has been tremendously refined, and it is too threatening to millions of English-Chinese translators and interpreters who are working on the translation industries, with the fare of losing jobs. The Law of Accelerating Returns, together with the Moore's Law, give out a signal that biological brains are much weaker, and AI is far more than superb. The application of exiting MT, like Iflytek, has achieved the multilingual instant translation, and it makes communication easier. Then think about those who work as translators or interpreters. What could they do to fight against MT? What can translators do to continue to demonstrate value in the translation industry?

One of the pioneers of Translation Studies (TS) is Holmes. In 1972, he outlined the scope of the relatively young area of translation studies. He demarcated the field into three broad areas: the descriptive branch where translations are described, the theoretical branch which explains the theories that inform the process of translation and the applied branch where the information gained from the former two branches is used for practical application. Holmes's concepts helped in the evolution of translation theories that concentrated more on the process of translation and the reception of the translated text in the target language was placing translation within specific contexts and attempting to look at them in a comprehensive manner.

Statement of the Problem

AI has the potential to significantly reshape the translation field, as shown in numerous aspects like speech recognition, image reading and information retrieval, etc. This paper focuses on the following three research questions:

- Will the English-Chinese translators and interpreters lose their job in the era of AI?
- what can translators do to fight back against Machine Translation?
- What can translators prepare for the era of Machine Translation and continue to demonstrate value in the translation industry?

Definition of Terms

Artificial Intelligence (AI): Ricardo (2013) cited Minsky's definition "Artificial Intelligence is the science of making machines do things that would require intelligence if done by men", and he put a stronger definition by replacing "intelligence" into "kinematics" (P247).

Machine Translation (MT): Various methodologies have been devised to automate the translation process (Tripathi & Sarkhel, 2010). It processes by means of rule-based, example-based, statistical, hybrid and neural MT.

Computer assisted translation (CAT): Huang, Zhao, Ma, Zhou & Zhang (2016) wrote, "CAT is a common way on language translation in which a human translator uses a software to perform and facilitate the translation process" (p.1).

Translation Studies (TS): James S. Holmes (1972) firstly put forward, "Translation studies is to be understood as a collective and inclusive designation for all research activities taking the phenomena of translating and translation as their basis or focus" (P 176). To make it simple, Traditional TS focuses on the interlingua translation between two different verbal sign systems (Munday, 2016).

Natural Language Processing (NLP): NLP is "a field of computer science, AI, and computational linguistics, concerning the interactions between computers and human natural languages", whose aim is "to discover effective theories and methods that enable humans to communicate better with computers using natural language", and it focuses on the actualization of natural language communication via computer system (Zhuang, Wu, Chen & Pan ,2017, p.5).

Purpose of the Study

This paper studies the MT and human transition, and it would encourage progress on extending and integrating the many disparate components of intelligent systems: reacting,

planning, learning, perception, and reasoning. It aims to seek a harmonious way for both parts in the era of AI.

Significance of the Study

The goal of this study is to formulate a set of challenge problems for the translation field. This paper will help English-Chinese translators and interpreters to eliminate the panic brought from the news or science magazines. In addition, it will give possibly constructive advices and strategies at the end.

Delimitation of the Study

The research conducted in and through the Karrmann Library at the University of Wisconsin-Platteville. Primary searches will be conducted via the Internet through EBSCO host with ERIC, Jstore, and Google/Google Scholar as the primary sources. Key search topics included “Artificial Intelligence”, “Machine Translation”, “Translation Studies” and “Neural Language Processing”.

Methodology

The research paper employed library research and focused on the most accurate and effective translation by virtue of the cooperation of human translation and machine translation. This project focused on the challenges and opportunities the English-Chinese translators and interpreters have, and finally it made conclusions and recommendations.

Chapter II Review of Literature

In this part, four perspectives will be reviewed in the order of "MT Background and Its Major Breakthrough", "Machine Translation & Computer-assisted Translation", "Human Translators' Threats & Weakness", "Human Translators' Strengths & Opportunities" and "Future of Human Translators". Jain, Kulkari & Shah depicted the relationship among deep learning and machine learning and AI. They (2018) wrote "Machine learning is a subset of AI", and "Deep learning and neural networks are subsets of machine learning" (P.161), which worked and aimed to analyze diverse types of data and improves machine knowledge. Their Natural Language Processing (NLP) studies involved "making computer systems to perform meaningful tasks with the natural and human understandable language" (P.161).

MT Background and Its Major Breakthrough

MT speech recognition and text processing have reached a new era in the history of interpreting and translation. Jain, Kulkari & Shah (2018) listed computational algorithms and approaches in text processing as below:

1. Long short term memory (LSTM), whose primary element model, is Recurrent Neural Network (RNN), which is chunk of neural network which can remember values. Other model has an extra state called peephole connection which lets the gate to check status of cell state and before dropping the data from network.
2. Compared with Statistical Machine Translation, Neural Machine Translation (NMT) is much more accurate in translation because it's capable to use algorithm to learn linguistic rules on its own from Statistical Models.

A recurrent neural network (RNN) is a class of artificial neural networks that were proposed in the 1980's (Rumelhart, Hinton, & Williams, 1988). Lai, Zhao & Bao confirmed his

model “adds the attention mechanism to the encoder-decoder structure, and the translation performance is greatly enhanced” (p.23).

Referring to Seq2Seq model, Jain et al. (2018) reviewed the two types of recurrent neural network, i.e., encoder network and decoder network, and technically demonstrated the translation procedure under this domain, but left no critical conclusion. Lai, Zhao & Bao (2018) argued the Seq2Seq model can guarantee that every input sentence has the same number of words and they concluded “decoding process can be simply understood as using a series of algorithms to return the word with the highest probability corresponding to the input vector to get the optimal output sequence” (P23).

Jain et al. (2018) also talked the limitations of NMT, and the dominant problems are prepositions. Then, the Voice Processing Algorithms are presented in acoustic modeling (contains the references of the individual sounds that make up a word) and Connectionist Temporal Classification (CTC). This part demonstrated the efficiency and accuracy of interpreting process; however, the disadvantages exist when corpus is missing in the database of written text and thus the algorithm would malfunction. Lai, Zhao & Bao composed experiments and conducted a result: the effect of Character segmentation on corpus is obviously higher than that of word segmentation on corpus (P.25).

Jain, Kulkari & Shah held positive attitude and suggested various applications would reduce the user's workload and time, giving appropriate and efficient output.

Machine Translation (MT) & Computer-assisted Translation (CAT)

For a better understanding of the essential principles underlying MT, Craciunescu, Gerding-Salas & Stringer-O'Keeffe (2004) looked into the functioning of the human brain, i.e., understanding words out of context, words in sentences and words in situations. They

highlighted five kinds of knowledge functioned during translation process: 1) source language; 2) target language; 3) the equivalents between the two languages; 4) the subject field as well as general knowledge, both of which aid comprehension of the source language text; 5) the customs and conventions of the source and target cultures. In comparison, MT is an autonomous operating system with strategies, i.e., direct strategy, the transfer strategy and the pivot language strategy.

MT can only explain the general meaning of the vocabulary, not the detailed interpretation and appropriate meaning. Besides, many MT resources do not provide the function of input/display of speech symbols, nor the function of translating from/to audio. Zhang (2017) developed a MT APP to bridge this gap by providing audio visual translation for better understanding of abstract meaning of words for people who do not understand the culture related to the language being translated. At the same year, Zhuang et al. wrote, "computer vision aims at simulating human perception capability with the power of computational modeling of the visual domain" (p.8). Computer vision investigates an effective visual computing pipeline that can carry out multiple "visual tasks ranging from low level (e.g., image feature analysis) to high level (e.g., visual recognition), resulting in the realization of image understanding in a human-like way" (Zhuang et al, 2017, p.8).

Craciunescu et al, (2004) made a distinction between MT and CAT which originate from different computational approaches. MT aims at "assembling all the information necessary for translation in one program" and has translation done without "human intervention". But automatic translation systems are not competent to produce immediate usable texts due to "different denotations and connotations of words and word combinations" of languages and they are dependent on context. That's why MT is limited to concrete situations and is recognized as

primary means of saving time, rather than a replacement for human activity. In contrast, CAT uses numerous tools to ensure an accurate and efficient translation work, e.g., terminology databases and translation memories, which is a needs-oriented approach. The process of CAT is more complicated than MT's, where "the computer becomes a workstation" and "the translator has access to a variety of texts, tools and programs: for example, monolingual and bilingual dictionaries, parallel texts, translated texts in a variety of sources and target languages, and terminology databases". Thus CAT "gives the translator on-the-spot flexibility and freedom of movement, together with immediate access to an astonishing range of up-to-date information. The result is an enormous saving of time".

Discussing on the domain of MT, William (2017) argued the usage of MT in medical field is valid for specific as well as general MT usage where judgement, empathy and creativity is involved, i.e., almost all field of life. He mentioned the artificial intelligence (AI) responsible for MT is susceptible to foibles and some AI's would tend to "loopthink" and in turn tend to utilitarian. He proposed that like humans in a group, there exist a way of thinking called "groupthink" which is "nondeliberate suppression of critical thoughts as a result of internalization of the group norm's". What similar is a "loopthink" for AI's where it resembles a stubborn person refuses to listen, by following a certain course of action and by rationalizing the information where unnecessary, e.g., a kinetic analogy would be the computer in a hypothetical self-driving automobile programmed to minimize the death toll in the event of unavoidable harm. In an inevitable crash situation, suppose that the car is headed straight toward a child with no time to stop. The human driver, intending to avoid hitting the child, swerves the steering wheel to the left, the car's computer, which ultimately commands the vehicle, disregards the instruction because it misinterprets as real a brightly-lit billboard on the left displaying pictures of three

people—or, perhaps, two people and an AI robot. Now in relation to translation, the importance correlation is that the MT done by AI is vulnerable to foibles, unnecessary rationale, immorality, stereotyping and censorship which needs to be addressed in order to make it more useful.

Human Translators' Threats & Weakness

Pan (2016) pointed out the future intelligent perception systems should not only simulate the brain effectively in terms of mechanism (referred to as 'human-like perception'), but also surpass the human brain in terms of performance (referred to as 'trans human perception'). This is one of the most noteworthy features of the next-generation AI, called 'AI 2.0'. Lumeras and Andy (n.d.) analyzed MT's strengths where human translators are less capable. They are conducted in four dimensions: terminology consistency, in glossary creation, spellchecking, preprocessing data, translation speed and translation scoring.

The Law of Accelerating Returns predicted the future will be far more surprising than most observers realize: few have truly internalized the implications of the fact that the rate of change itself is accelerating (Kurzweil, 2001). Machine Learning reforms in a stunning speed; however, our brain capacity is comparably limited. As the Chinese story goes, one Chinese emperor offers a reward to the inventor for his new beloved game (the game of Go), and the inventor asks for a single grain of rice on the first square, two on the second square, four on the third, and so on. The Emperor quickly granted this seemingly benign and humble request. As a result, the emperor went bankrupt as the 63 doublings ultimately totaled 18 million trillion grains of rice. At ten grains of rice per square inch, this required rice fields covering twice the surface area of the Earth, oceans included. Nowadays, the algorithms of AI machines are developing alike. This is the greatest threat to biological brains and individual translators and interpreters.

Human Translators' Strengths & Opportunities

Traditional deep learning models often rely on labeled data (Krizhevsk *et al.*, 2012), which is very difficult to obtain, thus human translators and interpreters have their advantages. Example of medical Chinese to English translation, as Lumeras and Andy put out, "低复发率" (lit. relapse level/rate low) MT encounters the difficulty in disentangling the right meaning; nevertheless, human translators come with natural saying "low risk of relapse". Human translators' rephrasing abilities are inherent. Another example was given, "复方药" (fu'fang'yao) in English is "a composed drug containing a mix of natural ingredients, such as herbs and powders used in Traditional Chinese Medicine". MT faces the limitation of rephrasing terms and expanding the translation. The future human translators should be bilingual and with PC developed competence in an integrated entity. Human translators are needed to construct bilingual corpus which builds up Artificial translation, because of the richness of the language, the integrity of the structure of the translation, the style of the translation, the appreciation of the word, the misunderstanding of the meaning of word meaning, etc. (Krizhevsk *et al.*, 2012).

. John Hutchins (2003) claimed "is unlikely to be a threat to the livelihood of professional translators" because human translators are more satisfactory and cost-effective(P,22). Referring to the spoken translation field, he argued interpreters of diplomatic and business, for example, are undeniably inevitable; however, MT systems are opening new areas where human translation has never featured, so there is a need and also a trend that Machine Translation and human should co-exist in harmony and without conflict.

MT helped human translators save amounts of time during the process of translating. The new working model transferred a human translator's role into a dictator who gave instructions and input corpus. Theoretically, it removes tedious tasks from human translators and free them

up to be increasingly creative. The newly developed Neural Machine Translation (NMT), an end-to-end learning approach for automated translation, is capable to sense translators and make both human and machine smarter (Wu et al., 2016). Zhao (2016) delineated that human translators are expected to build a stored bilingual corpus using computer skills, to apply technology to control the translation process of machine translations. Human translators focus on the accurate calibration after machine translation, and the process of checking translation loyalty is the process of demonstrating their own advantages.

Future of Human Translators

Green et al., (2015) has done a mindful discussion on natural language translation, MT and mixed approach. There has been significant development in human computer interaction (HCI) which is focused more on empirical approaches to usability and human factors, both of which generally aim to make machines more useful to humans. In practice, professional translators use suggestions from machine aids to construct final, high-quality translations. But how do we go beyond simple correction of machine mistakes? The corrections also help the machine, which can update its model to produce higher-quality suggestions in future sessions. In this positive feedback loop, both humans and machines benefit, but in complementary ways.

In 1951, Israeli philosopher Yehoshua Bar-Hillel observed the near-term achievement of "pure MT" was either unlikely or "achievable only at the price of inaccuracy." He then argued in favor of mixed MT, "a translation process in which a human brain intervenes." He then went on to define the now familiar terms pre-editor, for intervention prior to MT, and post-editor for intervention after MT.

Bar-Hillel's proposals were in the spirit of the more skeptical faction, which believed machine augmentation of existing human facilities was a more reasonable and achievable goal.

After years of research and development the finding that post-editing translation takes as long as manual translation is evidence of an interface problem. Surely even early MT systems generated some words and phrases correctly, especially for scientific text, which is often written in a formulaic and repetitive style. The question then becomes one of human-computer interactions: how best to show suggestions to the human user.

Recent Chinese-English translation research showed the importance of the conjugation and intergrowth of machine translator and human translator. Zhao (2016) gave three reasons on the topic:

1. Human translators and machine translators have the same existent condition composed by their similarities and logical relations;
2. The differences between human translators and machine translators are the complementary conditions for them;
3. The primitiveness of human translators and the development of machine translators leads to the relationship between prototype and model with them, and thus human translators have the characteristics of replacement and changeability.

Comparing a list of English to Chinese translation by MT and human translators, Zhao pointed out that MT can't compete with human translators and he focused on the language cultural connotation where MT fail to reach. In addition, semantic feature, ambiguity determination and cultural analysis of language makes professional translators unparallel. Nevertheless, MT has advantages and features which human translators are unable to surpass, they are: 1) Lower costs; 2) Shorter translation times; 3) Timely service; 4) Consistent translation; 5) High accuracy. Exactly, those are what anxieties and threats that AI has brought

in. Therefore, the future should conjugation and intergrowth of machine translator and human translator

Summary

The four domains of literature review present threats and weakness and study the strength and seek opportunities of human translators. It shows human translators face problems such as low working efficiency, higher possibilities in making general mistakes, unable to reform and keep pace with the industrial needs. In addition, researchers approved that human translators and interpreters have great opportunities performing better in the working field, as long as they manage to build stored corpus using computer skills and take part in controlling the MT process and checking the fidelity. Green S., Heer J., & Manning.D.C (2015) went through a set of experiments and made a conclusion that human-machine system designed for language translation benefits both human users--who produce higher--quality translations--and machine agents, which can refine their models given rich feedback.

MT is around us since 1950's, a lot of work has been done for pure MT which has proved to be unreasonable goal, and there is not a single reliable MT platform available till date. HCI model is being followed by the world, i.e., to say there is human interaction prior to and after MT; however, HCI model has its limitations as post-editing of the translation takes longer if not same than the conventional human translation. Although the HCI benefits from the correction provided by the users, it still needs years to improve.

Chapter III Conclusions and Recommendations

For years to come, human translator will be the core part of quality translation, and this role will not be taken away by MT. As for the scientific translation is concerned, MT is getting better, but still it is used as a preliminary low-cost tool. Without a human translator, it is not to be trusted as reliable translation. Human translators using MT as an initial guess often end up spending more time in correcting the grammatical / connotational mistakes, initial guess from MT might save time in the near future but perhaps at the cost of creativity, as languages evolve with time, so a general MT learning from past might never catches up with the able human translator; however, since scientific text is a different creed, it is anticipated that in near future translators might not have an edge over their MT counterparts.

This paper reviews the interaction and relations of new devices and systems. MT is a subset of AI, and it is a successful application of deep learning NLP. Deep learning algorithms analyze assorted data sets through algorithm again and again and improve the machine knowledge according to the output obtained. NLP is an integral area of computer science where machine learning is broadly applied. Machine learns the syntax and meaning of human language, process it and gives the output to users. As discussed, interactive machine translation reduces manual translation repetitive and mechanical operation. This field aims to make the human and computer interaction cushy but prompt.

As an English major Chinese student, I have witnessed how MT has been transferred the information and the pathway that English learner used to have. In Taiwan, where I had my exchange program, English instructors encourage students in using Trados (Google translation application), in which way the English to Chinese translation works are carried on faster and easier. It rapidly iterates repetitive parts of the design process, rather than a person manually

repeating it. Trados itself manipulates data to interrogate various concepts swiftly, cutting out a huge amount of repetitive work. In addition, with the unparallel machine learning ability, which could gain experience from the process and use that knowledge to improve efficiency and outcomes, Trados gets smarter and smarter as it accumulates larger volumes of past data, and it is personalized and nearly customized because of its “memory”. Thus, translator plays a role of a creative editor not merely a translator. This paper puts forward that there is no need to panic facing the challenges of the era of artificial intelligence. Instead, human translators need to proactively deal with its impact in various creative and technical ways.

To sum up, based on the analysis of research in the translation field, AI has immense potential to help address some of the biggest challenges that translators and interpreters face. At the meantime, AI systems would greatly benefit from advancements in theory, algorithms, and hardware to enable more robust, reliable, and intelligent perception. Human translators and interpreters are encouraged to enhance/ advance their computer skills and to co-work/ interact with the machine. Eventually, enjoy the benefit and convenience that the new era may bring in.

References

Bar‐Hillel, Y. (1951). The present state of research on mechanical translation. *Journal of the Association for Information Science and Technology*, *2*(4), 229-237

Bahdanau, Dzmitry, Kyunghyun. Cho, and Yoshua. Bengio (2014). Neural Machine Translation by Jointly Learning to Align and Translate. *Computer Science*.

Craciunescu, O., Gerding-Salas, C., & Stringer-O'Keeffe, S. (2004). Machine translation and computer-assisted translation. *Machine Translation and Computer-Assisted Translation*.

Caferra, R. (2013). Artificial Intelligence. In *Logic for Computer Science and Artificial Intelligence* (pp. 245-257). Hoboken, NJ, USA: John Wiley & Sons.

Cheshire JR, W. P. (2017). Loopthink: A Limitation of Medical Artificial Intelligence. *Ethics & Medicine*, *33*(1), 7.

Fiederer, R., & O'Brien, S. (2009). Quality and machine translation: A realistic objective. *The journal of Specialized translation*, *11*, 52-74.

Green S., Heer J., & Manning.D.C (2015). Natural Language Translation at the Intersection of AI and HCI. *Communication of the ACM, vol58.*

Holmes, J. S. (1972). The name and nature of translation studies. *Translated*, *2*, 67-80.

Hutchins, J. (2003). The development and use of machine translation systems and computer-based translation tools. *International journal of translation*,*15*(1), 5-26

Huang, G., Zhao, C., Ma, H., Zhou, Y., & Zhang, J. (2016, August). MinKSR: A Novel MT Evaluation Metric for Coordinating Human Translators with the CAT-Oriented Input Method. In *China Workshop on Machine Translation* (pp. 1-13). Springer, Singapore.

Jain, A., Kulkarni, G., & Shah, V. (2018). Natural language processing. *International Journal of Computer Sciences and Engineering* 6(1).

Kurzweil, R. (2001). The Law of Accelerating Returns, KurzweilAI. Net.

Krizhevsk, A., Sutskever, I., Hinton, G. (2012). Image Net classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, p.1097-1105.

Lumeras, M. A., & Way, A. (n.d.). On the Complementarity between Human Translators and Machine Translation. *HERMES-Journal of Language and Communication in Business*, (56), 21-42.

Linying, Z., Qiang, S (2017). 9. Research on English Machine Translation Based on the Basic Word Analysis. *Revista de la Facultad de Ingeniería*, *32*(16).

Lai, W., Zhao, X., & Bao, W. (2018). Tibetan-Chinese Neural Machine Translation based on Syllable Segmentation. *Technologies for MT of Low Resource Languages (LoResMT 2018)*, 21.

Munday, J. (2016). *Introducing translation studies: Theories and applications*. Routledge.

Pan, Y.H., (2016). Heading toward artificial intelligence 2.0. *Engineering*, 2(4):409-413.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5, 3.

Tripathi, S., & Sarkhel, J. K. (2010). Approaches to machine translation.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Klingner, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

Zhao LianBin. (2016). The connection with conjugation & intergrowth of machine translator and human translator. *Foreign Studies*, *4*(3), 100-103.

Zhuang, Y. T., Wu, F., Chen, C., & Pan, Y. H. (2017). Challenges and opportunities: from big data to knowledge in AI 2.0. *Frontiers of Information Technology & Electronic Engineering*, *18*(1), 3-14.