

Machine Translation and Chinese College Students' English Learning

Caiying Wang

School of Foreign Languages
Quanzhou Normal University
Quanzhou, Fujian, China 362000

Abstract—Nowadays, machine translation system has gradually gained popularity among college students, and it brings more and more benefits to them. This paper makes an analysis and comparison on lexical and syntax level of four machine translation systems commonly-used by Chinese college students. And it puts forward some relevant suggestions and methods to help students to use them and improve their English learning efficiency.

Keywords—machine translation system; Chinese college students; English learning; human translation

I. INTRODUCTION

Nowadays, Chinese modern higher education underwent radical changes due to the digital revolution in the past decades. The advent of digital or computer-mediated learning is not merely some kind of internal “experiment” within the broader educational environment. It is a social, economic, and cultural shift of startling proportions from which the higher learning cannot insulate itself. Under this circumstances, Ministry of Education of the People’s Republic of China issued the latest file in 2010, “The new teaching mode should be supported by modern information technology, especially by network technique. English learning and teaching can be free from time and place towards the development of the personality and independent learning.”

MT made strides in China in the past decades. Machine translation (hereinafter MT) system is a useful tool for college students’ English learning, that is to say, machine translation is conducive to language study to some extent. However, its inherent flaws, such as improper choice of words and inappropriate clause identification, together with students’ individual differences, make it unsuitable for all students; various MT systems such as Baidu and Tencent have been come up with by some large tech companies. Under these circumstances, this paper makes an analysis and comparison of the readability and acceptability of four available and most commonly-used machine translation systems in China and puts forward some relevant suggestions and methods to help students to use them and improve their English learning efficiency.

II. LITERATURE REVIEW

A. The Definition of Machine Translation

What’s machine translation? Daniel & Martin (2009) defined MT as an automatic process of transferring text from one human language to the target language by using context information.

B. The Available Chinese Machine Translation Software

The history of MT system is only more than one hundred years since German scholar W. Rieger first used the term—Machine Translation (MT) in 1903. While the history of MT in China can be traced back several decades ago. At present, Chinese college students have access to many available Chinese machine translation systems. According to the survey involving 1000 students (200 English majors and 800 non English majors) in Chinese universities showed that the commonly-used MT systems are as follows: Baidu Translate, Youdao Dict., Kingsoft PowerWord, and Google Translate.

C. Literature Review

MT plays an important role in college students’ English learning. Hence, there are few studies conducted to evaluate which MT software suits better for Chinese college students’ English learning on the ground of the text quality of MT. Previous research focused on MT approaches (Tripaths, 2010; AlAnsary S., 2011; Lawrence Muchemi et al., 2016), the quality of MT translation (Atef Odeh AbuSa’aleek, 2016), the ways to improve the quality of MT translation (Dimitar Shterionov, 2016; Ji Jing, 2017). Few research centered on MT and college students learning.

III. RESEARCH METHODOLOGY

A. Objectives of the Study

This study makes an attempt to learn about the performance of MT systems and tries to find out which system best suits Chinese students’ English learning based on their readability and acceptability on lexical and syntax level.

B. Questions of the Study

Are there any significant differences in the four MT outputs on lexical and syntax level and their accuracy?

C. Research Materials

We take the translation part of TEM8(Test for English Majors[China] Band 8) and the translation versions of the four MT systems as the research materials.

“生活像一杯红酒，热爱生活的人会从品出无穷的美妙。将它握在手中观察，它的暗红有血的感觉，那正是生命的痕迹。抿一口留在口中回味，它的甘甜有一丝苦涩，如人生一般复杂迷离。” (TEM 8, 2013) (Shēnghuó xiàng yībēi hóngjiǔ, rè'ài shēnghuó de rén huì cóngzhōng pǐn chū wúqióng dì měimào. Jiāng tā wò zài shǒuzhōng guānchá, tā de àn hóng yǒu xuè de gǎnjué, nà zhèng shì shēngmìng de hénjī. Mǐn yīkǒu liú zài kǒuzhōng huíwèi, tā de gāntián yǒu yīsī kǔsè, rú rénshēng yībān fùzá mílí.) (TEM 8, 2013)

- Baidu Translate(hereinafter A)¹: Life is like a cup of wine, the love life of the people will discover inexhaustible wonders from it. It will be in the hands of observation; it is blood red feeling, that is the trace of life. Take a sip of it, there is a bitter sweetness, such as complex life blurred.
- Youdao Dict.(hereinafter B)²: Life is like a glass of red wine, deeply loves the life will get the goods on the endless beauty. Will hold it for observation, its dark red with the blood of feeling, that's the trace of life. Sip what aftertaste, it is sweet to the soul have a bitterness, as complicated as life is blurred.
- Kingsoft Power Word(hereinafter C)³: Life is like a glass of red wine, love of life from the products of infinite wonderful. It will be held in hands to observe, feel it is dark red with blood, that is the trace of life. Close lightly a mouth in the mouth aftertaste, it sweet have a slightly bitter taste, such as general complex life blurred.
- Google Translate(hereinafter D)⁴: Life like a glass of red wine, people who love life will be out of infinite wonderful. Hold it in the hands of observation, it's dark red blood feeling, that is the traces of life. Sip a bit to stay in the mouth after taste, it's sweet and a bit bitter, such as the general complex blurred life.
- Reference Version: “Life is like a glass of wine; people who love it discover inexhaustible wonders from it. Hold in the hand and gaze at it, the dark red color is reminiscent of the blood, which is the impress of life. Take a sip of it and appreciate the taste, the bittersweet flavor is exactly the same with life, which is complicated and blurred.”

¹ A. <http://fanyi.baidu.com/>

² B. <http://fanyi.youdao.com/>

³ C. Kingsoft power word

⁴ D. <https://translate.google.cn/>

IV. ANALYSIS AND FINDINGS

A. Analysis

1)lexical errors

There are a lot of vocabularies in English and Chinese which have the same form but differ in part of speech and meaning. This difference is easier to understand and to be dealt with in human translation, but it is a big obstacle in the machine translation which often generates the wrong translation.

TABLE I. CHINESE CHARACTER

Chinese character	Baidu Translate	Youdao Dict	Kingsoft Powerword	Google Translate
红酒 (Hóngjiǔ)	wine	red wine	red wine	red wine
热爱生活的人 (Rè'ài shēnghuó de rén)	the love life of the people	deeply loves the life	love of life	people who love life
品出 (Pǐn chū)	discover	get	pretermisson	will be
抿 (Mǐn)	Take a sip	Sip	Close lightly a mouth	Sip
像 (Xiàng)	is like	is like	is like	like
无穷的 (Wúqióng de)	inexhaustible	endless	infinite	infinite

“红酒” (Hóngjiǔ): According to the definition given by Cambridge Dictionary, “red wine” means a wine with a dark red color that is made from black grapes, including their skins. While “wine” is defined as “an alcoholic drink that is usually made from grapes, but can also be made from other fruits or flowers. It is made by fermenting the fruit with water and sugar.” The first definition coincides with Chinese “红酒” (Hóngjiǔ).

“热爱生活的人” (Rè'ài shēnghuó de rén) is a noun phrase which can be translated into an attribute clause like Version D “people who love life”, the other three versions are wrong because of word-by-word translation.

“品出” (Pǐn chū) is a verb here, Version A is better than the other three versions, especially Version C omitted this word. The other two are mistranslated.

“抿” (Mǐn) is a verb here. Except that Version C omitted this word, the other three are right.

“像” (Xiàng) in Chinese original of this paragraph is a verb and can be translated into English “is like”. While “like” in English can be acted as a verb which equals Chinese meaning “喜欢” (Xǐhuān); “like” as a preposition in English is equivalent for “像……一样” (Xiàng...yíyàng) in Chinese. So Version D is wrong here.

“无穷的” (Wúqióng de) According to Collins Learner's English-Chinese Dictionary, “inexhaustible” means “that cannot be entirely consumed or used up”, “endless” means “tiresomely long; seemingly without end”, “infinite” means “having no limits or boundaries in time or space or extent or magnitude”. The most suitable one is “inexhaustible”.

2) *Syntax errors*: Like Chinese version, there are three sentences in all MT versions. On syntax level, A has three complete sentences. Although B, C, D are also comprised of three sentences, which involve the fragments of sentences, such as “Will hold it for observation” and “deeply loves the life” in B; “feel it is dark red with blood” in C and “that is the traces of life” in D. What is more, there are some grammar mistakes, such as “it sweet have” in C.

B. Findings

Compared with the above four versions of the same translation material translated by four commonly-used MT systems with the reference version of TEM 8 on morphology and syntax level, we can find that each MT system is far from able to produce error-free text. Errors and mistakes on lexical and syntactic level greatly affect intelligibility of MT outputs, because “syntactic errors irrevocably change the grammatical relations of sentence components, while lexical errors give rise to sentences full of confusing words. Short sentences had less syntactic errors, but suffered from wrong lexical selections.” (Yukiko Sasaki Alam, 2017)

Compared with the other three versions, the version of Google Translate is the most adequate and acceptable in regard to its adequacy and acceptability in the domains of words and syntax. “Most domestic MT products are younger than Google Translate, and they are not on the same level as Google in terms of algorithms. Therefore, it will be difficult for domestic companies to catch up with Google” said Zhang Xuetao, Director of the Language Service Competence Assessment and Training Center of TAC (Ji Jing, 2017).

V. SUGGESTIONS

The above analysis and findings shows that there is a critical need for further research of this area for the following reasons.

A. Choosing the Proper MT

From the aforementioned examples, each MT system has its pros and cons; with the development of science and technology, more and more MT software will emerge. Before college students turn to MT systems, they had better learn to select the appropriate an MT system for the specific language weighting on the analysis of a language resources, linguistic richness of the language versus the paradigm requirement.

B. Combination of Machine Translation and Human Translation

Although MT brings us convenience and high-efficiency, communication in human language is context-embedded. Due to the fact that source and target languages are usually quite different with respect to sentence (or text) construction, As Dimitar Shterionov (2016 : 32) pointed out “ Despite the advances in current SMT systems, their translation quality cannot be compared to human translation, and human post processing is often required.” Therefore, when college students do the translation works, it is necessary to have a good combination of machine translation and human translation to decide whether the words are understandable, the sentences are

grammatical, the text is logical and so on, which aims to improve the quality of translation. Some researchers reckon that human and machine translators should join hands because of the inadequacy of English translated by MT.

C. Using machine translation as a teaching tool

Although attention has been paid to machine translation's accuracy, Dr. Ana Nino of the University of Manchester has researched some of the advantages in utilizing machine translation in the classroom. One such pedagogical method is called using “MT as a Bad Model.” (Ana Nino, 2009). Because of the inherent flaws in words, grammar, sentences and discourses of MT, MT as a Bad Model gives a wide variety of teaching materials for teachers to apply in the classroom, such as word-reordering, subject-verb agreement and so on. For example, by analyzing the lexical errors, the teacher can guide the students to observe the semantic coherence of the relevant constituents; Many verbs have different meanings when used in different verb patterns or subcategorizations (Hunston, et al. 2000), and misuse of verbs and/or verb phrases can be frequently found in students English learning, it is widely known that MT systems s are easily accessible for teachers and students. Four different roles of MT in the language class have been identified: MT as a bad model, MT as a good model, MT as a vocational training tool (especially in the form of translation memories, pre-editing and post-editing), and MT as a “CALL tool” (Niño, 2009). MT system will have a profound influence on the teaching of Languages for Academic Purposes (Michael Grovesa, et al. 2015).

D. Using machine translation as a learning tool

MT systems have made promising progress and are becoming increasingly available for students' English learning. On one hand, some imperfect translations forces the language learners to identify inconsistencies or incorrect aspects of a translation to enhance college students' English learning which will help students sharpen their writing and translation skills; on the other hand, MT systems can be used to check for appropriate terminology and phraseology of specialized languages. The surveys conducted by some researchers (Belam 2003; Kliffer 2005) showed that post-editing serves as a complement to language learning.

VI. CONCLUSION

As discussed above, there are significant differences in four MT outputs on lexical and syntax level and their accuracy. What's more, current MT systems often leave some words untranslated which may cause problems because these untranslated words may be important to the underlying meaning of a sentence or even a paragraph. For instance, “回味” (Hu íwǎ) in the above-mentioned Chinese version is a pun which refers not only to appreciate the wine, but also to appreciate the life. However, Version A omits the translation of it; in Version B and C, it is translated into “aftertaste” and in Version D “after taste”, the “taste” of “回味” (Hu íwǎ) can't be fully conveyed by these translations. Moreover, some words including some cultural elements which do not have equivalents or have different meanings in English. Therefore, teachers should give the first priority to choose a most suitable

free online MT system when teaching and give some suggestions for students' learning.

Although MT still has some problems, it facilitates college students English learning. The functions such as the translation of webpages, software localization, code conversion, interlingual translation etc. provide great convenience when they process words, browse webs, read, write or translate. Snell Hornby(1988/1995) once said, "There is no doubt that the product of technology, no matter how complicated, can not be comparable to the creative ability of the human brain." Human and machine translators should join hands to derive a more fluent and accurate translation, especially for those complicated contexts or in difficult professional fields. What's more, it is of the importance to proofread and correct the MT-generated texts to make sure the fluency and adequacy of the translation.

Much emphasis was laid on the MT systems' practical value of using MT output for teaching and learning, the limitations and the inherent flaws of MT systems deserve much consideration, so both teachers and students should raise their awareness as to the complexity of translation and language learning. What we should do is to make an attempt to minimize the disadvantages, and to maximize the pros to enhance our teaching and students' learning.

REFERENCES

- [1] AlAnsary S. Interlingua-based Machine Translation Systems: UNL versus Other Interlinguas. In 11th International Conference on Language Engineering, Ain Shams University, Cairo, Egypt. 2011.
- [2] Alam Y. Knowledge Required for Avoiding Lexical Errors at Machine Translation[J]. International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 2017.
- [3] Atef Odeh AbuSa'aleek. The Adequacy and Acceptability of Machine Translation in Translating the Islamic Texts,J. International Journal of English Linguistics; Vol. 6, No. 3; 2016,PP.185-193.
- [4] Belam J. Buying up to falling down": A deductive approach to teaching post-editing[C]//MT Summit IX Workshop on Teaching Translation Technologies and Tools (T4)(Third Workshop on Teaching Machine Translation). 2003,PP. 1-10.
- [5] Daniel, J., & Martin, J. H. (2009). Speech and language processing: An introduction to natural language processing, speech recognition, and computational linguistics. Englewood Cliffs, N: Prentice-Hall.
- [6] Dimitar Shterionov. Machine transion with brains,J. MultiLingual. July/August, 2016, pp.31-34.
- [7] Groves M, Mundt K. Friend or foe? Google Translate in language for academic purposes[J]. English for Specific Purposes, 2015, 37PP. 112-121.
- [8] Hunston S, Francis G. Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English[J]. Computational Linguistics, 2000, 27(2).
- [9] Ji Jing. Human and Machine Translators Join Hands[N]. Beijing Review 2017 , 60 (4)
- [10] Kliffer M D. An experiment in MT post-editing by a class of intermediate/advanced French majors[J]. Practical applications of machine translation, 2005,PP.160-165.
- [11] Lawrence Muchemi, Wanjiku Nganga, Benson Kituku. A review on machine translation approaches,J. Indonesian Journal of Electrical Engineering and Computer Science. Vol. 1, No. 1, January 2016, pp. 182 - 190.
- [12] Niño A. Machine translation in foreign language learning: language learners' and tutors' perceptions of its advantages and disadvantages[J]. ReCALL, 2009, 21(02) PP.241-258.
- [13] Snell Hornby, Mary. Translation Studies: An Intergrated Approach.M. Amsterdam & Philadelphia:John Benjamins Publishing Company, 1988/1995.
- [14] Tripaths,Sarkhel. K. Approaches to machine translations. Annals of Library and information studies.2010, 57. pp. 388-393.