

MACHINE TRANSLATION EVALUATION IN SNS IN TERMS OF USER-CENTERED ORIENTATION

Kim Euna

Department of English Language & Literature, Busan, South Korea

ABSTRACT

This Study explores the role of machine translation by creating a corpus of text from the one of SNS, Facebook, and analyzing and evaluating the corpus data in terms of User-Centered Translation (UCT). For the data to examine, Reuter's Facebook account with language pair of English and Korean was selected due to the fact that the posts are open to the public and use a formal structure of sentences. Based on the corpus, a questionnaire was made to actually see the response from users who are following the Reuter's account and using translation function.

KEYWORDS

Machine Translation, Social Network Services, Corpus, User-Centered Translation, Target Reader

1. INTRODUCTION

Machine Translation (MT) has been applied to many potential areas such as translation of web page or translation of Social Network Services (Wang, 2007). The research about MT has been widely studied but it is still considered to have limits in terms of translation quality (Wilks, 2009). However, the output of MT has shown nearly zero errors compared to human translation since it has been improved from phrase-based, statistical based to the neural based algorithm. The success of MT has brought the promising communication tool in SNS where users around the world share their stories in their own various languages.

In view of the fact that machine translation errors are nearly ignorable, then the users of translation should be taken into account to evaluate the result of MT. But whether it is statistical or neural data, information on translation users are not included or translation process does not have revision process to consider users' information. Therefore, the objectives of this research is to design a corpus with the posts uploaded in Facebook, evaluate and analyse the translation quality along with the translation users' information to see whether the result of translation corresponds to the purpose of users and to help users adapt to the function of MT so that they can gain better communication efficiency.

2. BACKGROUND

2.1. Development of Machine Translation

A number of translation systems is developed and the well-known translation includes rule-based machine translation (RBMT), statistical machine translation (SMT), example-based machine translation (EBMT), knowledge-based machine translation (KBMT), hybrid machine translation (Shi, 2013) and recently neural machine translation (NMT). Lastly, Facebook announced that they already have adopted NMT for their translation function.

2.2. Neural Machine Translation System

The main issue on the MT was the time delay of the translation process and translation quality. A number of methods have been proposed to solve this problem and the introduction of neural-network-based MT systems showed a great improvement in simultaneous translation and accuracy of standard (Gu et al. 2017). The NMT was proposed by (Kalchbrenner & Blunsom, 2013), (Stuskever et al., 2014) and (Cho et al., 2014a). The basic idea of NMT is to encode a different length of sentences into a fixed length vector that would summarize the whole sentence, then, decode this encoded vector in the target language to achieve the translation of source sentence (Verma, 2017).

3. DATA USED

3.1. Corpus

For the data to examine, the posts of the Reuter's Facebook account with language pair of English and Korean were selected due to following reasons; 1) public account, therefore, open to any users 2) official account using a formal structure of sentences for the post. Then, a bilingual corpus of 4000 posts was created and evaluated for the quality of the translation. Each post has a picture with one short English explanation with a length of a sentence, hash tags, which are the feature of the SNS, date, and the writer's name of the post.

3.2. Translation users

There are more than 1.2 million followers of the Reuter's account and among them are the translation users whose mother tongue is not English. To distribute questionnaires to Korean users who follow the Reuter's account, the number of users is not enough to observe the result and to lead to the general fact. Therefore, 20 people were randomly chosen and asked them to follow the Reuter's account, use the translation function for the Reuter's posts according to their purposes and distributed the questionnaires.

4. DISCUSSION

4.1. Data evaluation

The corpus data was first evaluated to analyse the quality of translation by three different researchers and they identified the reasons of decreasing the quality of translation. Various types of automatic evaluation metrics have been developed such as BLEU (BiLingual Evaluation Understudy), WER (Word Error Rate), PER (Position-independent word Error Rate), and TER (Translation Error Rate) (Mohamed, 2012). But those metrics require references to calculate the

translation quality, three parameters were chosen as a criteria in this research to evaluate the quality of translation; 1) understandability, 2) grammaticality, 3) lexical correctness. The understandability means the amount of information that readers are able to understand, the grammaticality means to find out the grammatically ill-formed sentences and the lexical correctness means to identify the untranslated words and suitability of chosen words in the given context (Bohan et al. 2000). Then, 3-point value scale was given to each sentence such as bad, neither bad nor good, and good and computed the average across subject area.

4.1.1. Evaluation result

With the language pair of English and Korea, 4000 sentences or posts were evaluated and their average score of translation was calculated. For the understandability, 2465 sentences were rated as good, 1321 as neither bad nor good, 214 as bad. For the grammaticality 2189 sentences were rated as good, 1008 as neither bad nor good, 803 as bad. For the lexical correctness, 2532 sentences were rated as good, 1503 as neither bad nor good, 145 as bad. Every sentence was evaluated by three persons to reduce subjectivity.

4.2. Questionnaire to translation users

The questionnaire was distributed to the translation users after they followed the Reuter's account and used the translation function. The questions include 1) users information (age, occupation, fluency of English) 2) purpose of using English Reuter's account (Education, Information gathering, translation, using as secondary source, etc.) 3) evaluation of translation quality in 3 scales with 12 examples 4) phenomena causing the bad translation quality (with open-ended question) 5) satisfaction level of translation function in Facebook in accordance with their purposes.

4.2.1. Discussion

The sample group of 20 people was distributed with the questionnaire of 17 questions, which are 16 closed-ended questions and one open-ended question. The result of the questionnaire is shown in Table 1, Figure 1 and Figure 2.

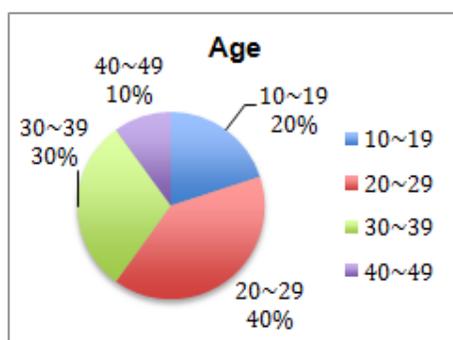


Figure 1. Age of sample group

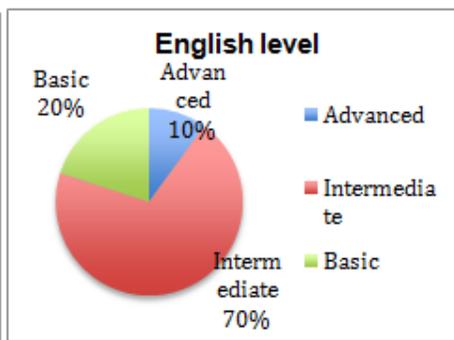


Figure 2. English level of sample group

Table 1. Result of the questionnaire

Age	10-19	4
	20-29	8
	30-39	6
	40-49	2
	Total	20 people
Occupation	Student, professors, engineer, government officer, no job, etc.	
	Advanced	2
	Intermediate	14
	Basic	4
	Total	20 people
Evaluation of translation quality (12 sentences)	Good	3
	Neither bad nor good	10
	Bad	7
	Total	12 sentences
Phenomena causing the bad translation quality	Unsuitable lexical choice, idioms, unknown word, untranslated word, relative clauses, completeness of sentences, word order, adverbs, etc.	
Satisfaction level	Extremely satisfied	2
	Satisfied	1
	Neither satisfied nor dissatisfied	6
	Dissatisfied	11
	Extremely dissatisfied	0
	Total	20 people

People who participated in this research ranged from 10 to 49 years old and their occupation varies such as a student, professor, engineer, government officer or no job at all. Their English level, which was rated by themselves showed that two people were in advanced, 14 were in intermediate and four were in the basic level. Among 12 sentences chosen for the questions, each four sentence was chosen by the research according to the average scale of their translation quality, which was rated by three researchers. Three over 12 sentences were rated as good, 10 as neither bad nor good, and 7 as bad. For the open-ended question, 20 people answered that the phenomena causing the bad translation quality were the unsuitable lexical choice, idioms, unknown word, untranslated word, relative clauses, completeness of sentence, word order, adverbs, etc. Final question result showed that two of 20 people were extremely satisfied with the translation function of the Reuter's posts, one was satisfied, six were neither satisfied nor dissatisfied and 11 were dissatisfied.

5. CONCLUSIONS

Although the development of the TM has proven its perfect translation result, the application of the TM in Facebook showed some significant errors that affected the satisfaction level of users. This is because the algorithm of the TM has been developed but the translation process or the memory data does not include the users' information, which means the final result does not correspond to the users' purpose and does not reach to the level satisfying the users. Therefore, they still require the revision process performed by professional translator or users themselves to guarantee the translation quality. This research aims to evaluate the translation quality of the MT applied to the Facebook, but also to help users to adapt to the MT function so that they can achieve their purposes of using SNS and to get the better quality of translated texts. In future research, a retranslation of each sentence by human will be carried out and will be used as a

reference, therefore, evaluation system such as BLEU, WER, PER, and TER could be used to get a more objective score of the translation quality.

ACKNOWLEDGEMENTS

The author would like to convey thanks to Prof. Yoon Hwayoung, her supervisor who always gives valuable suggestions and guidance for completion of her research.

REFERENCES

- [1] Wilk, Yorick (2009) Machine Translation, Its Scope and Limits, Springer.
- [2] Wang, HaiFeng(2007) "Pivot Language Approach for Phrase-Based Statistical Machine Translation", Machine Translation, Vol. 21, No. 3, pp165-181.
- [3] Shi, Chun Qi (2013) "User-Centered Design of Translation Systems", Doctor Thesis Series of Ishida & Matsubara Laboratory, Kyoto University.
- [4] Gu, Jiatao, Neubig, Graham, Cho, Kyung Hyun & Li, Victor O.K. (2017) "Learning to Translate in Real-Time with Neural Machine Translation", Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, Vol. 1, pp1053-1062.
- [5] Verma, Ajay Anand & Pushpak, Bhattacharyya (2017) "Literature Survey: Neural Machine Translation", Conference proceedings.
- [6] Mohamed, Amine Cheragui (2012) "Theoretical Overview of Machine Translation", Proceedings ICWIT 2012, pp160-169.
- [7] Bohan, Niamh, Breidt, Elisabeth & Volk, Martin (2000) "Evaluating Translation Quality as Input to Product Development", Proceedings of LREC 2000.

AUTHORS

Kim Euna is currently a representative PhD candidate in Pusan National University in Korea

