HOW ACCURATE CAN YOU GO?
AN INVESTIGATION TOWARD FACEBOOK'S MACHINE TRANSLATOR

Fatimatuz Zahro[1], Irham[2]

Department of English Literature, UIN Maulana Malik Ibrahim

Jalan Gajayana No. 50 Malang, 65144, Indonesia

Email: (1) zahrafu14@gmail.com (2) irhamaladist@gmail.com

Abstract

This study attempts to carry out meticulous investigation upon the translation errors by machine translator in Facebook. Having known that machine translator in Facebook often performs inaccurate translation, we tried to classify its translation errors in the first place. We focused on *Bahasa* Indonesia to English translation, and then providing logical explanation of the phenomena. The data of this study were collected from several users' posts on Facebook that are translated from *Bahasa* Indonesia into English. We consider the frequency and content analysis, the textual, grammatical and cohesiveness of translation. The result revealed 55 errors within 18 Facebook posts. Most errors were classified into the syntactic errors with 22 or 40.01% errors in total. Meanwhile, the mechanic errors were found 16 items or 29.09% and the semantic errors consist of 23.63%. The rest of the data were categorized as the cultural errors which consist of 4 issues or 7.27%. This study, afterwards, may help the readers understand the common translation mistakes that the robots produced. Therefore, the society will be more aware and the possibility of misunderstanding may decrease. Besides, it will also be helpful for future social media developers to consider more about the language features on their applications.

Keywords: error analysis, facebook, machine translator, social media, translation

1. INTRODUCTION

A language is a system of arbitrary vocal symbols by means of which a social group cooperates (Bloch &Trager, 1942, p. 5) between or among one another. Language also plays a pivotal role in society in a way that it is used as a medium of communication. By so doing, people can keep interacting and transferring the idea. People speaking the same language may not have problem to understand the message being delivered. Nevertheless, those speakers of different language may encounter difficulties. Here is why translation is indisputably needed, especially when people intend to communicate with other societies from different language in virtual way. In recapitulated understanding, translation is the transfer of thoughts and ideas from one language to another (Brislin, 1976, p. 1).

The super-fast development of social media makes world-wide communication almost unavoidable. It enables people to interact with others living in another part of the world, eating different foods, and speaking different languages. As the biggest social platform, Facebook is currently giving its users an easier way to communicate internationally by providing the machine translator. This feature automatically translates the posts on the users' timeline that help them understand those posts. However, translation cannot be separated from some aspects: the

textual, grammatical and many others. Those terms are important to make our words understandable to be read in another language.

For the above mentioned reason, errors in translation are often produced by Facebook's machine, for example when translating *bahasa* Indonesia into English. Unlike the other translation devices such as Google Translate, Facebook's machine translator has a bigger possibility to make mistakes. Meanwhile, a half of 2.23 billion Facebook users see the translation results monthly. Therefore, we think this problem should be responded.

Studies on machine translation have been quite many in number, for instance, Wongranu (2017) on Thai-English translation, discussing the types and causes of errors. In addition, Károly (2011) studied English-Hungarian translation errors in terms of specialized EU genre translation by using the textual analysis and retrospective interviews. Yousofi (2014) , on the other hand, investigated translation errors of Iranian novice English translator. The study focused on the translation results of Iranian graduate and undergraduate students which resulted in the errors of three areas of language: linguistics, cultural and stylistics

Some scholars also put more attention on how machine translation performs and affect readers. For example, Fedrico, Negri, Bentivogli, and Turchi (2014) examined the effects of machine translation upon human perception. They employ mix-effect models to formulate a robust evaluation. In addition, Irvine, Morgan, Carpuat, Daume, and Munteanu (2013) tried to perform the analysis on machine translation errors within macro and micro level of the analysis. While these studies were too focus on machine translation and its impact, the present study aims to investigate syntactic errors performed by Facebook machine translation. It is worthwhile to do so in a way that linguistic aspects in translation may affect the resulted meaning(s).

Literature Review

According to Dulay, Burt, and Karshen (1982, p. 1), error analysis aims to find the flawed side of learner speech or writing. They are those parts of conversation or composition that deviate from some selected norm or mature language performance. Error analysis consists of the description, explanation and identification of the errors represented in learner language (Ellis & Barkhuizen, 2005, p. 51). Sunardi (2002) in Tandikombong, Atmowardoyo, and Weda (2016, p. 4) stated that error analysis is an activity to reveal errors found in writing and speaking. This concept is also used to correct the errors and improve the translation competence.

This study is benefited from several existing theories in both translation and syntax. We accounted Pojprasat's (2007) work on translation errors in which syntactic, semantic and cultural errors were then claimed to be important aspects in doing the analysis. The syntactic errors are the mistakes found within the sentence structure. We therefore make use of Chomsky (2015) and Sag, Bender, and Wasow's (2001) classifications in identifying possible syntactic errors. These are omission, tense, word order, addition, voice, agreement, preposition, double verb, and part of speech.

In regard with the syntactic errors, Dulay et al. (pp. 138-139) classified it into four types: omission, addition, misformation and misordering. We also based the present analysis on Wongranu's (2017) work who studied the analysis of translation errors in Thai to English translation. He categorized the syntactic error categorization into 20 aspects including the countability, determiners, tenses, agreement and fragments. The theory used was basically referred to the level of correctness in the textual, referential and cohesive level triggered by Newmark (1987, pp. 22-23). The two other aspects analysed with the theory proposed by Pojprasat (2007) are the semantic and cultural errors. The semantic errors are defined as the mistranslation of words that includes a single word, collocations or

idioms. Meanwhile, the cultural errors are any errors that are produced by the cultural differences. However, we would not only focus on the three aspects but adding the analysis the mechanic errors which include the punctuation and capitalization as well.

Since the study is focused on the automatic translation, we did not make use of English students' work as the objects as other previous studies. Instead, we paid more attention to the robot translator that is currently popular within society. The analysis would cover the syntactic, semantic, mechanic and cultural errors.

2. DATA COLLECTION AND METHODOLOGY

The data source of this study is the posts made by 15 Facebook users. There are 19 Facebook posts that have been translated from *Bahasa* Indonesia to English. Since the technology of artificial intelligence has never been able to do it, the machine translator still cannot identify and translate abbreviation, informal and slang words until today. Therefore, the researcher only takes the posts with formal language as the data, thus each word of the sentence will be translated.

We employed both qualitative and quantitative research design. The qualitative research was aimed to bring a detailed account of one or more cases (Johnson & Christensen, 2000, p. 46). Besides, the quantitative research was intended to explain the phenomena analysed with a numeral (Aliaga & Gunderson, 2002, p. 151). In terms of data analysis processes, we identified errors in Facebook's translated posts and classified them into 4 categories: syntactic, semantic, mechanic and cultural errors. We then did performed a statistic upon the observed phenomena and elaborate correspondingly.

3. RESULTS AND DISCUSSION

The classification of translation errors produced by Facebook's machine translator was summarized in Table 1. The classifications were broken down into 4 parts: syntactic, semantic, mechanics and cultural errors.

Syntactic Errors

Syntactic errors held the highest percentage as can be seen in Table 1. It means the machine translator was more likely to produce errors that were related to the structure or grammar of the translated sentence. There were some syntactic error types as mentioned in Table 2.

Table 1 Total Percentage of Errors

Classification of Error	Frequency	Percentage
1. Syntactic errors	22	40.01
2. Mechanic errors	16	29.09
3. Semantic errors	13	23.63
4. Cultural errors	4	7.27
Total	55	100.00

The most frequent syntactic errors were the omission, tense and word order. The omission errors found in this machine translator's production were commonly happening within the omission of determiner and preposition.

As can be seen in the first example, the word *"criminal"* did not have a determiner, thus this translation included as an omission error. The second example shows an omission of the preposition *"on"* that should be placed after the verb *"focus"*.

Table 2 Percentage of Syntactic Errors Categorization

Category of Error	Frequency	Percentage
1. Omission	4	20.00
2. Tense	3	15.00
3. Word order	3	15.00
4. Addition	2	10.00
5. Voice	2	10.00
6. Agreement	2	10.00
7. Preposition	1	5.00
8. Possessive	1	5.00
9. Double verb	1	5.00
10. Part of speech	1	5.00

Other types of error that mostly found were the tense and word order errors. The tense errors were found in the use of past and present tense. Facebook's translator often had no time signal agreement in translating a sentence. For instance, in translating this sentence: *"...sudah sepatutnya doa tidak dijadikan sebagai sekedar rencana cadangan",* Facebook used past tense as it is translated: *"...that it was supposed to not be made as a backup plan",* whilethere was no past time signal. This analysis was also supported by its previous sentence that is translated into present tense. Besides, there were also other errors such as the omission of *"doa"* and word order. Hence, the correct translation will be: *"...that prayer is not supposed to be made as a backup plan".* Another word order error was produced in translating a phrase such as *"aktivitas membacakan buku"* which was translated into *"activities reading books".* The correct translation of this phrase would be: *"the book reading activity"* or by making it simpler: *"the reading activity".* This phrase also contaminated with a countability error where Facebook could not differentiate the plural and singular form. The word *"buku"* was only mentioned once, thus it was supposed to be singular, but Facebook added *–s* and made it plural.

Table 3 Examples of Omission Errors

Sentence	Facebook's translation	Correct translation
1. Kriminal sadar tindakannya salah.	Criminal has his actions wrong.	A criminal realizes that his actions are wrong.
2. ...saya mesti focus nulis buku.	....I should focus writing books.	...I should focus on writing books.

Mechanic Errors

The mechanic errors held the second-lowest percentage. These mechanic errors included the errors of punctuation and capitalization. As can be seen in Table 5, Facebook's translator often made some mistakes in term of mechanics. Unlike Google Translate, this translator turned some capitalized words that used to be a name into uncapitalized words and vice versa.

Table 4 Examples of Mechanic Errors

Category of Error	Word/Phrase	Facebook's translation
1. Capitalization	JaBoDeTaBek	Jabodetabek
	...yang menerbitkan buku.	...Who published a book.
2. Punctuation	...dibuat dalam bentuk gambar. Inilah hasilnya.	...made in the form of This is the result.

In the first example of capitalization, the word *"JaBoDeTaBek"* was a name, thus it should be translated in the same way by capitalizing some letters. However, Facebook's translator did not do it that way, not even capitalize the first letter. The second example shows how Facebook's translator capitalized a word that was not supposed to be capitalized. Whilst in the example of punctuation error, Facebook's translator omitted the period punctuation and directly turned to the next sentence.

Semantic Errors

The semantic error ranked the third highest percentage. As we know, the meaning is the most influential thing in translation. However, Facebook's translator seemed to ignore this aspect too much as the examples of error provided in Table 4.

Most of the semantic errors happened within important words that were missing. In the first example, the translator omitted one important word: *"penjual"*. Since this word became the modifier of the word *"yellow rice"*, it is influential to make the context correct and not ambiguous. Besides, this kind of error was too obvious since *"yellow rice"* is a name of an inanimate object, thus it is illogic that *"his mother used to be a yellow rice"*. Those examples show that the omission error is important to be noted since it will also impact the meaning of the sentence or the semantic aspect. Unfortunately, Facebook's translator still ignored such errors.

Table 5 Examples of Semantic Errors

Sentence/Phrase	Facebook's Translation	Correct Translation
1. Ibunya dulu penjual nasi kuning...	His mother used to be a *yellow rice...*	His mother used to be a *yellow rice seller...*
2. Siapa yang paling berperan dalam prosesmu menjadi penulis?	Who is *most role* in your process to be a writer?	Who is *the most important/influential* in your process to be a writer?
3. Sila-sila pancasila	*Please-please* pancasila	Pancasila*principles*

As illustrated in the other examples, Facebook's translator did a mistake by translating the word *"paling berperan"* into *"most role"*, while it did not make any sense. However, the word *"most important"* or *"most influential"* would be more appropriate to be used in this sentence. In the last example, Facebook's translator might assume that the word *"sila-sila"* refers to the Malayan word *"sila"* or *"silahkan"* in *Bahasa* Indonesia, thus it translated the word into *"please-please"*. This mistake is also what the researcher found in Google Translate. Both Facebook's translator and Google Translate had not been able to identify the word *"sila"* in Indonesia which means *"principle"*. This result proved that even Facebook's translator is run by a robot,the vocabulary knowledge should be questioned. In addition, it often gives inappropriate word choice too.

Cultural Errors

As what mentioned previously, cultural errors are related to different culture that affects the translation result. There were some terms in *Bahasa* Indonesia which are absence in English. Hence, those words are difficult to be translated. There are 4 issues detected within the data as follows:

Table 6 Examples of Cultural Errors

Original sentence	Facebook's translation
1. Lagi latah, pamer dikit...	Again *flinching,* show off a little...
2. Kok beli makan aja?	Why buy *eat*?

In the first example, the word *latah* was translated to *flinching,* whereas actually, both words have different meanings. In *Bahasa* Indonesia, *latah* in such contextmeans following the trend of someone else. In another context, *latah* can be defined as mocking. Meanwhile, *flinch* means making a fast movement as a response of surprising thing. Therefore, the translation made by the machine translator is not relevant at all.

Moving to the next example, the word *makan* was translated to *eat,* whereas *makan* in this context is not a verb but a noun instead. In English, the word *eat* is an activity of consuming something, while the object consumed is called *food.* On the other hand, in Bahasa Indonesia, the word *makan* can be either the activity or the object as like *minum*or*drink*. Therefore, the translation result should be: *why do you buy foods only?*

4.   CONCLUSION

The findings found that the translation errors produced by Facebook's machine translator consist of syntactic, semantic, mechanic, and cultural errors. The syntactic error held the highest percentage of the errors, it is 40.01%. The semantic error placed the second highest percentage, it is 29.09%, followed by the mechanic error with 23,63% percentage and the cultural error with 7,27% total of errors.

According to the data, analysis and findings explained previously, some suggestions can be given for the Facebook's users and developers. For the users, they should be more aware of Facebook's translation result. This awareness is important since there have been so many mistakes the machine translation makes that, of course, will lead to a misunderstanding between users at worst. Another suggestion is given to the social media developer. As their productions will be used by society, they hold a big expectation of society on how their machines will bring a better system, especially in term of language. As one of the communication supporter, Language holds a big

role in social media. Hence, making a language system that is effective and appropriate is very important to build a good connection between the users.

REFERENCES

Aliaga, M. & Gunderson, B. (1999).*Interactive Statistics.*New Jersey: Prentice Hall.

Bloch, Bernard &Trager, George L. (1942).*Outline of Linguistic Analysis.*Baltimore: Waverly Press.

Brislin, Richard W. (1976). *Translation: Applications and Research.* New York: GradnerPress Inc.

Chomsky, N. (2015). *Aspects of the theory of syntax*. Cambridge, Massachusetts: The MIT Press.

Dulay, H., Burt, M., and Krashen, S. (1982).*Language Two.*Oxford: Oxford University Press.

Ellis, R. &Barkhuizen, G. (2005).*Analysing Learner Language.* Oxford: Oxford University Press.

Federico, M., Negri, M., Bentivogli. B., and Turchi, M. (2014). Assessing the Impact of Translation Errors on Machine Translator Quality with Mixed-effects Models. *Empirical Methods in Natural Language Processing.* 1643-1653.

Irvine, A., Morgan, J., Carpuat, M., III, Hal Daumé, and Munteanu, D. (2013). Measuring Machine Translation Errors in New Domains. *Transactions of the Association for Computational Linguistics. 1*(1), 429-440.

Johnson, B., Christensen, L. B. (2000). *Educational Research: Quantitative and Qualitative Approaches.*Boston: Allyn & Bacon.

Károly, Adrienn. (2011). Translation competence and translation performance: Lexical, syntactic and textual patterns in student translations of a specialized EU genre. *English for Specific Purposes, 31*(1)*,* 36-46. doi: 10.1016/j.esp.2011.05.005

Newmark, Peter. (1987). *A Textbook of Translation.* New Jersey: Prentice Hall.

Pojprasat, Somboon. (2007). An Analysis of Translation Errors Made by Mattayomsuksa 6 Students. Retrieved from http://thesis.swu.ac.th/swuthesis/Eng(M.A.)/Somboon_P.pdf

Sag, I. A., Bender, E. M., & Wasow, T. (2006). *Syntactic theory: A formal introduction*. Stanford, Calif: CSLI.

Tandikombong, M., Atmowardoyo, H. and Weda, S. (2016).Grammatical Errors in the English Translation Made by the Students of English Study Program of UKI Toraja.*ELT Worldwide, 3*(1), 1-15.

Wongranu, Pattanapong. (2017). Errors in translation made by English major students: A study on types and causes. *Kasetsart Journal of Social Sciences, 38*(2)*,* 117-122. doi: 10.106/j.kjss.2016.11.003

Wulandari, Indra. (2014). Grammatical Errors Found in Articles' Abstracts of Indonesian Scholarly Journals.*Journal of English Education and Linguistics Studies, 1*(1), 12-30.

Yousofi, Nouroddin. (2014). Describing the Errors in the Translations of Iranian novice English Translators. *Procedia-Social and Behavioral Sciences, 98,* 1952-1958. doi: 10.1016/j.sbspro.2014.03.628

Yule, George. (2010). *The Study of Language.* Cambridge: Cambridge University Press.

Appendix 1: Data Source of the Study

No.	Username	Date and Time of Post	Original sentence *(Translated sentence)*
1	Andre Raditya	8 May 2018 02:05	Bagi mereka yang tahu hakikat perannya sebagai hamba.. Pasti akan menyadari, bahwa sudah sepatutnya doa tidak dijadikan sebagai sekedar rencana cadangan. *(For those who know his role as a slave.. Will surely realize, that it was supposed to not be made as a backup plan)*
2	Betty	8 January 2017 18:27	Mabuk resi... Huhu.... *(Drunk receipt... Huhu....)*
3	BintangNurul	10 May 2018 20:27	...Jadi kita disuruh bikin tarian mancanegara buat pelajaran SBD. Hukumnya sunnah sih. *(So we're supposed to make foreign dances for sbd. Law, though.)*
4	Boy Candra	9 May 2018 09.49	Entah kenapa, saya selalu merasa memiliki energy tambahan setiap ada penulis dari Sumatera Barat yang menerbitkan buku. Apalag ijika yang menerbitkan buku adalah anak muda dan senang mengembangkan diri mereka. *(I don't know why, I've always felt an additional energy there was a writer from West Sumatra Who published a book Especially if publishing books is young and happy to develop themselves.)*
5	Boy Candra	10 May 2018 19:31	Siapa yang patah hati boleh mengutuk dan menghukum orang yang membuatnya patah hati. *(Who breaks the heart can condemn and punish the people who break his heart.)*
6	Fayanna Ailisha Davianny	10 May 2018 20:51	Bulan April-Mei, SiDU mengunjungi 100 sekolah di wilayah JaBoDeTaBek, untuk mensosialisasikan program inidan membagikan bukuMenulis Bersama SiDU untukmelatih kemampuan menulis. *(In April-may, sidu visited 100 schools in the jabodetabek, to socialize this programe and share a writing book with sidu to practice writing skills.)*
7	Kharisma Maheswari	31 July 2018 13:08	Kok beli makan aja? Gak beli minum? *(Egads, why buy eat? Don't buy a drink?)*
8	Malkan Junaidi	10 May 2018 06:30	Jika kita baca sila-sila pancasila, adakah yang bertentangan dengan ajaran islam?

			(If we read please-please pancasila, did anyone contradict the Islamic teachings?)
9	Maya Lestari Gf	8 May 2018 15:09	...Kebanyakan audiens merasa heran karena bagi mereka, aktivitas membacakan buku seharusnya berlaku untuk anak yang belum bisa membaca. *(Most audiences feel surprised because of them, activities reading books should apply to children who can't read.)*
10	Maya Lestari Gf	9 May 2018 22:53	...Nah, karena Mahreen belum bisa menulis, maka plot ceritanya dibuat dalam bentuk gambar. Inilah hasilnya. *(Well, because mahreen hasn't been able to write, then plot the story is made in the form of This is the result.)*
11	Muhsin Labib	10 May 2018 17:46	Sayangnya, makna terorisme dibatasi pada tindakan. Padahal ia bisa dideteksi dari sikap intoleran dan hate speech berbungkus agama dan mazhab. *(Unfortunately, the meaning of terrorism is limited to action. He may be able to be able to be detected from intolerant and hate speech and madhhab)*
12	Muhsin Labib	10 May 2018 01:27	Kriminal sadar tindakannya salah. Teroris yakin tindakannya benar. Manakah yang lebih berbahaya? *(Criminal has his actions wrong. Terrorists believe his actions are true. Which is more dangerous?)*
13	Muhsin Labib	1 January 2018	Orang-orang yang mabuk ketenaran mengira banyak orang iri kepada ketenarannya, padahal yang "gilo" lebih banyak. *(People who are drunk of fame think a lot of people are jealous of his fame when the "Gilo" is more.)*
14	Noor H. Dee	10 May 2018 08:46	Sekarang, istilah "pengalihan isu" sering dipakai oleh mereka yang sebenarnya tidak tahu apa-apa tapi ingin terlihat tahu segalanya. *(Now, the term "Diversionary issue" is often used by those who don't really know anything but want to know everything.)*
15	Penerbit Indiva	7 May 2018 08.38	Siapa yang paling berperan dalam prosesmu menjadi penulis? *(Who is most role in your process to be a writer?)*
16	Sherina	23 November 2016 16:35	Lagi latah, pamer dikit, biar dikata norak no problemo. Gue terima SPP nih! Ucapkan, Alhamdulillah *(Again flinching, show off a little, let's say tacky noproblemo. I accept the tuition! Say, alhamdulillah)*

17	Siradj Kumar (in BubuhanUrang Samarinda (BURSAM))	16 April 2018 13:57	Muhammad Jusuf Hamka, Putra asli Samarinda, Ibunya dulu penjual nasi kuning di samarinda. *(Muhammad J. Hamka, the original son of samarinda, his mother used to be a yellow rice in samarinda.)*
18	Triana Retno Adiastuti	9 May 2018 07:20	Ada kalanya saya mestifokus nulis buku. Nggak bisa lirik kiri kanan, bahkan terpaksa nolak order ngedit dan nutup toko online buats ementara. *(Sometimes I should focus writing books. Can't make the right lyrics, even forced to take an address and close the store for a while.)*
19	Wa Ode WulanRatna	22 January 2018 06:44	Potret Bre Redana di Jurnal Ruang oleh Ade Irwansyah kali ini menarikdisimak bagi yang selalu penasaran dengan sosoknya. *(Potrait of BreRedana at JurnalRuang by Ade Irwansyah this time it is interesting to be curious about him.)*