GOOGLE NMT AND YANDEX NMT DIFFERENCES IN CLARITY AND NATURALNESS OF TRANSLATING ENGLISH TEXT INTO INDONESIAN: A COMPARATIVE STUDY ON MACHINE TRANSLATIONS

Anggun Citra Sasmi[1]*, Baharuddin[2], Lalu Ali Wardana[3]

[1][2][3] English Education Department, Faculty of Teacher Training and Education, University of Mataram, Indonesia

*Corresponding Author: angguncitra473@gmail.com

Abstract: Neural Machine Translation (NMT) has emerged as a transformative technology, revolutionizing the way we interact and communicate in a multilingual world. This study presents a comparative analysis of two prominent NMT systems, Google Translate (GNMT) and Yandex Translate (YNMT), focusing on their respective capabilities in translating English text into Indonesian while evaluating the clarity and naturalness of the translations. Qualitative descriptive methods is applied in this study. The analysis reveals that GNMT exhibits a superior level of naturalness and clarity over YNMT based on the analysis of three types of texts, procedural text, narrative text and descriptive text. Notably, both NMTs demonstrate enhanced proficiency in translating the texts, yet GNMT consistently maintains its edge over YNMT in terms of translation quality. Overall, this comparative study serves as a valuable contribution to the field of machine translation, providing valuable insights into the performance variations of GNMT and YNMT in the context of translating English into Indonesian. It is hoped that this research will stimulate further investigations, ultimately propelling the domain of NMT towards more accurate, clear, and natural translations across languages.

Keywords: *Neural Machine Translation, Google Translate, Yandex Translate, clarity, naturalness*

INTRODUCTION

The process of transferring a text's meaning from one language to another is called translation. Venuti (1995) defined translation as the process of swapping out the chain of signifiers in a text written in a foreign language with a chain of signifiers in the target language that the translator provides depending on his interpretation. Now humans can translate with the help of machines. Translating by machine translation is very convenient, the result is fast and the cost is low. In addition, machine translation can now be downloaded to smartphones, making it easy to translate words anytime, anywhere without having to carry a traditional dictionary.Machine translation or well known as neural machine translation (NMT). According to Wu et al. (2016), neural machine translation (NMT) is already widely used in industries such as Google, Microsoft, Facebook, Amazon, SDL, and Yandex. As mentioned earlier, Google

and Yandex are the two most widely used machine translations in the world. translation of the world. Google Translate: Google released an OT tool Google Translate (GT) in 2006. The tool consists of a phrase-based algorithmic translation model that "analyzes word pairs based on their frequency of use in a huge digitized dataset" (King, 2019, p. 2). Meanwhile, Yandex Translate: Yandex translate (formerly Yandex translation) is a website published by the Russian company Yandex for translating the text into another language. There are certainly many factors that determine the quality of translation as a product. According to Nababan et al. (2012), the quality of translation can be measured using the factors of translation accuracy, clarity or intelligibility, and naturalness or readability. The translator's main job is to convey the meaning of the source text to the target text. Translators sometimes simply translate text word for word, regardless of the difference in "style" between the two languages. Due to this phenomenon, The researchers are keen to conduct research focused on comparing Google NMT and Yandex NMT to translate English text into Indonesian with an emphasis on clarity and also naturalness. Besides that, the writer dedicated this study to facilitate others researchers, the students of the university, and everyone that has the same interest.

RESEARCH METHODS

Qualitative descriptive research is used in this study. The researcher's data collection includes three English texts: a procedural text titled "How to Make Scrambled Eggs," a narrative text titled "The Ant and the Dove," and a descriptive piece titled "My Father Bought Me A Puppy". Researchers employed text and data sheets as research tools to gather information to address research issues. To collect the data, this study used three steps: input and retrieval, testing and documentation. To analyze the data, this study used four steps: identification, classification and description. In addition, this study needs a rater or an assessor which is Mr. Dr. Baharuddin, M. Hum who is indeed qualified in the field of translation as an assessor to assess the quality of clarity and naturalness form each text translated by both NMTs.

FINDINGS AND DISCUSSION

The research findings and analysis of the data from the Indonesian translation of English texts are presented in this chapter. This is based on research by Larson, who claims that translation activity is commonly defined as the act of converting a source language (SL) into a target language (TL) in his book "Meaning-Based Translation" (1998: 531–533).

Table I. Percentage results of clarity and naturalness from the translated text by GNMT and YNMT

Text	Clarity		Naturalness	
	GNMT	YNMT	GNMT	YNMT
Procedure Text	65,62%	62,50%	50,00%	46,87%
Narrative Text	56,81%	54,54%	47,72%	50,00%
Descriptive Text	61,36%	59,09%	43,18%	43,18%

The table above presented comparison results from each translated text by GNMT and YNMT. The findings indicate that the highest percentage of translation was achieved by Google NMT either clarity or naturalness criteria. Furthermore, the data provides a detailed breakdown and the complete comparison results of clarity and naturalness analysis from each text.

Clarity Analysis in Procedure Text

The table below presents from 8 sentences, 5 sentences from procedure text translated by GNMT that were categorized as clear and 3 categorized as less clear. On the other hand, 5 sentences translated by YNMT were categorized as clear and 2 categorized as less clear also 1 categorised as unclear. The total score obtained by sum up each score from each sentence.

Table II. Comparison of clarity from translated procedure text by GNMT and YNMT

Criteria	Sentences	
	GNMT	YNMT
Very Clear	0	0
Clear	5	5
Less Clear	3	2
Unclear	0	1
Total scores	21	20
Percentage	65,62%	62,50%

Here is an example of Less Clear by GNMT while Unclear by YNMT:

The sentence in the procedure text:

ST: Repeat the previous step until all the eggs are <u>mostly cooked</u>.
GNMT: Ulangi langkah sebelumnya sampai semua telur <u>hampir matang</u>.
YNMT: Ulangi langkah sebelumnya sampai semua telur <u>sebagian besar matang</u>.

If the word ***"mostly cooked"*** is translated word for word then it means it is ***"sebagian besar matang".*** But in the context of what Indonesians are discussing it is not common to use the word ***"sebagian besar matang"*** to describe mostly cooked food but an ***"hampir matang"***. This synced with the work produced by GNMT that said ***"ulangi langlah sebelumnya sampai semua telur hampir matang"***. Therefore, GNMT would be categorized as less clear in this context. But on the other hand, YNMT translates it ***"ulangi langlah sebelumnya sampai semua telur sebagian besar matang"*** where it is not commonly used among Indonesians, so this context makes it categorized as unclear. On the other hand, why does both NMT could not categorized as very clear, since both NMT results show there's still ambiguity in part of ***"semua telur",*** like it still sound awkward for indonesian.

Clarity analysis in Narrative Text

The table below presents from 11 sentences, 4 sentences from narrative text translated by GNMT that were categorized as clear, 6 categorized as less clear and 1 categorized as unclear. On the other hand, 2 sentences translated by YNMT were categorized as clear, 7 categorized as less clear and 2 categorised as unclear. The total score obtained by sum up each scores from each sentences.

Table III. Comparison of clarity from the translated narrative text by GNMT and YNMT

Criteria	Sentences	
	GNMT	YNMT
Very Clear	0	0
Clear	4	2

Less Clear	6	7
Unclear	1	2
Total scores	25	24
Percentage	56,81%	54,54%

1. Here is an example of Clear by GNMT while Less clear by YNMT

The sentence in the narrative text:

ST: After walking around for a moment, she came to a spring.

GNMT: Setelah berjalan-jalan sejenak, dia <u>sampai</u> di sebuah mata air.

YNMT: Setelah berjalan-jalan sebentar, dia <u>datang</u> ke mata air.

If translated into word for word then ***"she came to a spring"*** in this context must mean ***"dia sampai atau tiba di sebuah mata air"*** because the context is that the ant was looking for a spring and then after walking around she reached a spring. This is so similar to the GNMT that ***"setelah berjalan-jalan sejenak dia sampai di sebuah mata air"*** where the GNMT can be described as clear. But on the other hand, YNMT describes it as ***"setelah berjalan-jalan sebentar dia datang ke mata air"*** which had nothing to do with the context being discussed because it was impossible for the Indonesian people to search for water but instead to use the word ***"dia datang ke mata air"*** where he was supposed to use the word ***"ketika dia sedang mencari air kemudian dia tiba atau sampai di sebuah mata air"*** therefore YNMT are categorized as less clear.

2. Here is an example of less clear by GNMT while unclear by YNMT

The sentence in the narrative text:

ST: <u>Guessing what he should do, the ant quickly bit him on the heel</u>.

GNMT: <u>Menebak apa yang harus dia lakukan</u>, semut dengan cepat menggigit tumitnya.

YNMT: Menebak apa yang harus dia lakukan, dan dengan cepat menggigitnya.

If this word is translated word for word then it means **"menebak apa yang harus dilakukan".** As we can see, the two translation machines translate this way. But actually, in the contextual of what is appropriate is **"memikirkan apa yang harus dilakukan".** In this we can see that from both machines the translators did not understand it that way. But how to determine that the GNMT is less clear while the YNMT is unclear is that we see from the sequel. Which **"the ant quickly bit him on the heel"** should mean is **"semut dengan cepat menggigit tumitnya"** and the translation is the same as the one on GNMT. On the other hand, the YNMT reads **"dan dengan cepat menggigitnya"** as we see that there is a section that is not translated **"the ant"** so that we can categorize that here YNMT is unclear.

Clarity analysis in Descriptive Text

The table below presents from 11 sentences, 5 sentences from descriptive text translated by GNMT that were categorized as clear, 6 categorized as less clear. On the other hand, 5 sentences translated by YNMT were categorized as clear, 5 categorized as less clear and 1 categorised as unclear. The total score obtained by sum up each score from each sentence.

Table IV. Comparison of clarity from the translated descriptive text by GNMT and YNMT

Criteria	Sentences	
	GNMT	YNMT
Very Clear	0	0
Clear	5	5
Less Clear	6	5
Unclear	0	1
Total scores	27	26
Percentage	61,36%	59,09%

Here is an example of less Clear by GNMT while unclear by YNMT

The sentence in the descriptive text:

ST: But Dad <u>can't see</u> her, so Dad decides to bring him home and <u>give him some comfort.</u>

GNMT: Tapi Ayah <u>tidak bisa melihatnya</u>, jadi Ayah memutuskan untuk membawanya pulang dan <u>menghiburnya.</u>

YNMT: Tapi Ayah <u>tidak bisa melihatnya</u>, jadi Ayah memutuskan untuk membawanya pulang dan <u>memberinya kenyamanan</u>.

The word ***"give him some comfort",*** If translated into Indonesian then the meaning is ***"memberinya kenyamanan".*** But in this context the word memberinya kenyamanan unusual used by the Indonesian people but will use words like ***"menghiburnya".*** Which is the same as the translation of GNMT when GNMT translated it with ***"ayah memutuskan untuk membawanya pulang dan menghiburnya".*** Therefore, here GNMT could be categorized as less clear. But on the other hand, YNMT could just translate it into ***"ayah memutuskan untuk membawanya pulang dan memberinya kenyamanan"***. Therefore, the YNMT is categorized as unclear. Moreover, both NMT could not categorized as clear even very clear since in the word ***"tidak bisa melihatnya"*** *from* ***"can't see her"*** means ambigue just sounds like the father is blind when he actually not. Actually, in this context, ***"see"*** should mean ***"menemukan"*** instead of ***"melihat"***.

Naturalness Analysis in Procedure Text

The table below presents 8 sentences from procedure text translated by GNMT, which 2 were categorized as Natural, 4 were categorized as Less Natural and 2 were categorized as Unnatural. On the other hand, 2 sentences translated by YNMT were categorized as Natural, 3 were categorized as Less Natural and 3 were categorized as Unnatural.

Table V. Comparison of naturalness from translated procedure text by GNMT and YNMT

Criteria	Sentences	
	GNMT	YNMT
Very Natural	0	0

Natural	2	2
Less Natural	4	3
Unnatural	2	3
Total scores	16	15
Percentage	50,00%	46,87%

Here is an example of Less Natural by GNMT while Unnatural by YNMT in procedure text:

ST: Repeat the previous step until all the eggs are <u>mostly cooked</u>.
GNMT: Ulangi langkah sebelumnya sampai semua telur <u>hampir matang</u>.
YNMT: Ulangi langkah sebelumnya sampai semua telur <u>sebagian besar matang</u>.

If this word is translated into word for word then it means it is *"sebagian besar matang".* But in the context of what Indonesians are discussing it is not natural to use the word *"sebagian besar matang"* to describe a mostly cooked food but *"hampir matang".* This synced with the work produced by GNMT that said *"ulangi langkah sebelumnya sampai semua telur hampir matang"*. Therefore, GNMT would be categorized as less natural in this context. But on the other hand, YNMT translates it as *"ulangi langkah sebelumnya sampai semua telur sebagian besar matang,"* where it is not commonly used among Indonesians, so this context makes it sounds unnatural.

Naturalness Analysis in Narrative Text

The table below presents 11 sentences from narrative text translated by GNMT, which 1 was categorized as Natural, 8 were categorized as Less Natural and 2 were categorized as Unnatural. On the other hand, 2 sentences translated by YNMT were categorized as Natural, 7 were categorized as Less Natural and 2 were categorized as Unnatural.

Table VI. Comparison of naturalness from the translated narrative text by GNMT and YNMT

Criteria	Sentences	
	GNMT	YNMT
Very Natural	0	0
Natural	1	2
Less Natural	8	7
Unnatural	2	2
Total scores	21	22
Percentage	47,72%	50,00%

1. Here is an example of Less Natural by GNMT while Unnatural by YNMT in the narrative text:

> ST: After walking around for a moment, she came to a spring.
> GNMT: Setelah berjalan-jalan sejenak, dia sampai di sebuah mata air.
> YNMT: Setelah berjalan-jalan sebentar, dia datang ke mata air.

If translated into word for word then *"she came to a spring"* in this context must mean *"dia sampai atau tiba di sebuah mata air"* because the context is that the ant was looking for a spring and then after walking around she reached a spring. This is so similar to the GNMT that *"setelah berjalan-jalan sejenak dia sampai di sebuah mata air"* where the GNMT can be described as less natural. But on the other hand, YNMT describes it as *"setelah berjalan-jalan sebentar dia datang ke mata air"* which had nothing to do with the context being discussed because it was impossible for the Indonesian people to search for water but instead to use the word *"dia datang ke mata air"* where he was supposed to use the word *"ketika dia sedang mencari air kemudian dia tiba atau sampai di sebuah mata air"* therefore YNMT are categorized as unnatural. However, both NMT could not categorized as very natural since there are still ambiguity in their translation results.

2. Here is an example of Less Natural by GNMT while Natural by YNMT in the narrative text:

ST: Seeing that the ant was in trouble, the dove quickly put off a leaf from a tree and dropped it immediately into the water near the <u>struggling ant.</u>

> *GNMT: Melihat semut dalam kesulitan, burung merpati dengan cepat melepaskan sehelai daun dari pohon dan segera menjatuhkannya ke air di dekat <u>semut yang meronta.</u>*
>
> *YNMT: Melihat semut dalam kesulitan, burung merpati dengan cepat melepaskan sehelai daun dari pohon dan segera menjatuhkannya ke dalam air di dekat <u>semut yang sedang berjuang.</u>*

If translated into word for word then ***"struggling ant"*** in this context must mean ***"semut yang sedang berjuang"*** because the context is that the ant was struggled to get out from the water. This is so similar to the YNMT that ***"semut yang sedang berjuang"*** where the GNMT can be described as natural. But on the other hand, GNMT describes it as ***"semut yang meronta"*** which actually different from the actual meaing in the context of the story. Therefore, GNMT are categorized as Less Natural. However, both NMT could not categorized as very natural since there are still ambiguity in their translation results.

Naturalness Analysis in Descriptive Text

The table below presents 11 sentences from narrative text translated by GNMT, which 1 was categorized as Natural, 6 were categorized as Less Natural and 4 were categorized as Unnatural. On the other hand, 1 sentence translated by YNMT were categorized as Natural, 6 were categorized as Less Natural and 4 were categorized as Unnatural. Therefore, In this text the results of both NMT is similar.

Table VII. Comparison of naturalness from the translated descriptive text by GNMT and YNMT

Criteria	Sentences	
	GNMT	YNMT
Very Natural	0	0
Natural	1	1
Less Natural	6	6
Unnatural	4	4
Total scores	19	19

Percentage	43,18%	43,18%

CONCLUSION

According to the translation result by GNMT and YNMT, in all types of text the researcher could not find either very clear and very natural sentences that are translated by GNMT and YNMT. Therefore, we can conclude that both NMTs are good at translating English text into Indonesian for some particular text in this study yet could not be very accurate as usual Indonesian itself. However, based on the results and discussion, it is obvious that GNMT opposed to YNMT, provides a clearer and more natural translation of English text into Indonesian. For more detail, GNMT is good at translating all types of text, but better at translating Procedure Text with a percentage of 65,62% in clarity and 50,00% in naturalness, while YNMT is good at translating Narrative Text with a percentage of 54,54% in clarity and 50,00% in naturalness. To sum up, from the whole results GNMT is still better in use than YNMT.

REFERENCES

Venuti, L. (2017). *The translator's invisibility: A history of translation*. Routledge.

Wu, Y., et al. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

King, K. M. (2019). Can Google Translate be taught to translate literature? A case for humanists to collaborate in the future of machine translation. *Translation Review*, *105*(1), 76-92.

Nababan, M., & Nuraeni, A. (2012). Pengembangan model penilaian kualitas terjemahan.

Larson, M.L. (1998). *Meaning-Based Translation*. University Press of America. (pp. 531).