

ERRORS IN LEXIS MADE BY THE MACHINE TRANSLATION FROM THE CAPTIONS OF NATIONAL GEOGRAPHIC INDONESIA'S INSTAGRAM

Ni Made Eni Kesumawati

Universitas Mahasaraswati Denpasar

enikesuma29@gmail.com

Ni Made Verayanti Utami

Universitas Mahasaraswati Denpasar

I Dewa Ayu Devi Maharani Santika

Universitas Mahasaraswati Denpasar

ABSTRACT

Many social medias provide machine translation to help the users understand any caption made in another languages. But it is something gives inaccurate translation which makes the meaning is not delivered properly. The study aims to identify types of semantic errors in lexis made by Instagram Machine Translation found in the National Geographic Indonesia's Instagram captions. Observation method and note taking technique is used in to collect the data from National Geographic Indonesia Instagram Account from May until October. This study used a descriptive qualitative method in analyzing types of semantic errors in lexis and applying the theory types of semantic errors in lexis by James (1998). The result shows that the confusion of sense relations in choosing a wrong synonym is mostly found in data source. This phenomenon occurred because Instagram Machine Translation literally translated the source language to the target language which have several synonyms and different meaning in each use without considering the overall context of the caption.

Keywords: semantic errors in lexis, Instagram, machine translation

INTRODUCTION

Semantics has a close relationship to the translation. Semantics allude to a study of meaning used to comprehend human expression through language. According to Katz (1972: 1), "Semantics is a subfield of linguistics represent considerable authority in the study of meaning". It manages sentences and other semantic rules with practically no plan of the syntactic or pronunciation. Consequently, semantics assumes a huge part as the study of meaning, innate at the degree of words, expressions, sentences and bigger units. The utilization of semantics in each language is undeniable in the translation process.

According to Larson (1998:3), "translation is basically an alteration of form". It shows that the meaning has been transferred from SL to TL with practically identical importance or semantic design. Budiman & Saifullah (2020) explain translation process involves the transferring of ideas in the source language to the target language. Semantics plays a critical part in the translation process,

guaranteeing that the importance of the SL is precisely sent and abstaining from misconception or disarray in the TL. Moreover, every language has a particular grammar, which may be hard for an interpreter to comprehend the importance of SL. This peculiarity frequently happens in the translation process.

As technology develops, creating machine translations for translating different languages simplifies the translation process. The users are able to enter source language and select the target language. Next, create a translation that requires machine translation. Therefore, machine translation quickly used to transfer the meaning of SL into TL. For example, Instagram machine translation. Instagram users take advantage of the translation engine available on Instagram to translate text or captions owned by the official accounts they follow. However, the users should have consideration for the lack of accuracy of the translation results generated by the Instagram engine. For instance, the translation of the caption on the National Geographic Indonesia Instagram account.

National Geographic Indonesia Instagram account always posts pictures and videos relates to science, nature and culture along with captions in Indonesia. Instagram provides an automatic machine translation for non-Indonesian people. Indeed, the translation results obtained by Instagram's machine translation are inaccurate and contain many errors. For instance, errors made by National Geographic Instagram captions on 31st July 2021, "... tempat itu juga menjual surplus barang-barang berharga terjangkau dan menyiapkan *makanan siap saji* yang sehat." translated into English by Instagram Machine Translation to be, "...it also sells surplus of affordable valuables and prepares healthy *ready food*." The term "*makanan siap saji*" is translated into "*ready food*" in target language. The term "*makanan siap saji*" in the source language is tied together and generates a meaning as a food that is served quickly and can be taken away to be eaten in the street. However, the collocation of "*ready food*" is unjustified and really makes no sense. Instagram machine translation translated the collocation "*makanan siap saji*" literally from source language into "*ready food*" in the target language. The suitable collocation for "*makanan siap saji*" is "*fast food*".

Using the found example, it was proved that there were found semantic errors in lexis in the process of translating the captions on Instagram of National Geographic Indonesia. This study is very useful in improving readers knowledge and considerations when translating one language to another using machine translation available on social media platforms like Instagram. Users should pay more attention to the results of the translation and mistakes made by Instagram's machine translation. This research focused on semantic errors in lexis by finding the types namely confusion of sense relation and collocational errors made by Instagram machine translation on the National Geographic Indonesia Instagram Captions.

DATA SOURCE

The data for this research were taken from National Geographic Indonesia Instagram captions. The National Geographic Indonesia Instagram account was selected as the data source because the captions made by National Geographic Indonesia's Instagram Machine Translation has 238 data containing semantic errors. National Geographic Indonesian Instagram accounts use Indonesian

language. Thus, many users, followers and viewers may not understand Indonesian very well. This research has obtained the latest data of posts from May to October 2021. In addition, the selected data were used as a sample to represent the entire data.

METHODS

The method applied in this study is descriptive qualitative method. Qualitative method is the method that applied to describe in detail the data and compiled systematically to obtain the comprehensive result. There were four steps in collecting and analyzing the data namely, Searching on National Geographic Indonesia's Instagram account, reading captions, taking a note of the sentences which containing semantic errors in lexis, classifying the data and analyzing the data based on the theory of James. C (1998) in his book entitled "Errors in Language Learning and Use: Exploring Error Analysis".

THEORETICAL REVIEW

Shormani & Al-Sobani (2012) define semantic errors are the offense the arrangement of English in Semantic rules. The rules in semantics are closely related to the word choice system. Moreover, the discrepancy between words and its meaning will generate the existence of violation. According to James (1998) there are two main types of semantics error in lexis namely, confusion of sense relations and collocational errors. confusion of sense relations can be divided into utilizing a generic term for the specific term utilizing a very specific term for the generic term, utilizing the inappropriate of two co-hyponyms, utilizing the inappropriate a pair of close-synonyms.

1. Confusion of Sense Relations

Lexicologists define lexicon represent the relationship among words and also its meaning. Lexical fields are a type of lexicosemantic clustering. As a result, it is fair to try to classify lexis errors using these methods. The most common errors consist of:

- a. Utilizing a generic term for the specific term such as superonym for hyponym which causes the meaning specifications
- b. Utilizing a very specific term for the generic term (hyponym for superonym)
- c. Utilizing the inappropriate of two co-hyponyms
- d. Utilizing the inappropriate a pair of close-synonyms

2. Collocational Errors

Collocations are the series of words or terms are frequently associated with. A psychologist's test for determining word-association relies on collocation awareness, especially as the focus is on integrative associations (rather than paradigmatic connections). A collocation is a gathering of at least two words that are quite often assembled to make a particular meaning. If an error in combining words in a collocation will affect the meaning conveyed. Especially to native speakers it will sound

very strange and violate the rules of proper language in collocation. In fact, listeners will identify oddities in the utterances conveyed when using collocation errors.

RESULTS AND DISCUSSION

Table 1. The Types of Semantic Errors in Lexis in the Translation of National Geographic Indonesia’s Instagram Captions

NO	Types of Semantic Errors in Lexis	Occurrence
I	Confusion of Sense Relations	
1	Utilizing a generic term for the specific term	19
2	Utilizing a very specific term for the generic term	10
3	Utilizing the inappropriate of two co-hyponyms	5
4	Utilizing the inappropriate a pair of close-synonym	136
II	Collocational Errors	68
	TOTAL	238

Based on the data table, the amount of semantic errors in lexis found in the National Geographic Instagram Captions are 238 data. From 238 data, there are 170 (71.42%) data of confusion of sense relations; a) 19 data (7.98%) of utilizing a generic term for the specific term, b) 10 data (4.20%) of utilizing a very specific term for the generic term, c) 5 data (2.10%) of utilizing the inappropriate of two co-hyponyms, and d) 136 (57.14%) of utilizing the inappropriate a pair of close-synonym. There are 68 (28.57%) of collocational errors. The most frequently found data is the utilize of the inappropriate a pair of close-synonym. This phenomenon occurred because Instagram Machine Translation used an inappropriate translation procedure by choosing the lexis in the target language literally through a term that can have several synonyms which have different meaning in each use without considering the overall context of the caption. The least founded data is utilizing the inappropriate of two co-hyponyms with 5 data (2.10%). This phenomenon occurred because the data that consist of co-hyponyms are rarely found on Instagram machine translation and usually the data have been translated correctly.

Types of Semantic Errors in Lexis

According to James (1988), semantic errors in lexis are divided into confusion sense relations and collocational errors.

1. Confusion of Sense Relations

There are four types of confusion of sense relations namely:

1.1 Utilizing a generic term for the specific term

In the process of translating the source language using terms that have specific meanings. If translated into the target language there are also specific terms that are commensurate with the source language. However, when the source language is translated

into the target language using terms that are too general, it will lead to the confusion for readers and it is difficult to specify the meaning of the source language. This error is classified as semantic errors in the use of general terms for specific terms that affect the readers understanding of the translation results.

Data:



Figure 1.1 Natgeo Indonesia's Post on 28th of September 2021

SL: Aroma manis itu menuntun kami pada rumah warga lokal yang membuat *gula aren*.

TL: The sweet smell led us to a local home making *sugar*.

(<https://www.instagram.com/natgeoindonesia/> on 28th of September 2021)

Semantic errors in lexis are found in the translation above and can be classified as the confusion of sense relations in choosing a general term for the specific term. The source language utilized a specific term. However, the target language utilized a general term in transferring the meaning of the source language. According to KBBI online, the term “gula aren” in the source language means a type of sugar made of liquid secreted from the flowers of trees of the palm family, such as coconut, palm, and siwalan. However, according to Oxford Learner’s Dictionary 2021, the “sugar” refers to a sweet substance, often in the form of white or brown crystals, made from the juices of various plants, used in cooking or to make tea, coffee, etc. sweeter. The term “gula aren” itself is a type of traditional sugar that has existed for a long time because it is made of palm tree sap which is processed traditionally. The term “gula aren” refers to a specific term which can be included as part of the term “sugar”. Therefore, the Instagram machine translation has failed to convey the specific meaning of the term “gula aren”. Instagram machine translation translated the term “gula aren” generally into “sugar”. The term “sugar” in the target language simply refers to a general explanation of the source language because there are many types of sugar such as coconut sugar, palm sugar, sanding sugar, coarse sugar, granulated sugar, etc. However, the term “gula aren” has specific meaning as only a kind of traditional sugar made of the nectar of the palm tree or *Arenga pinnata*. The Instagram machine translation is able to convey the meaning of source language correctly if

translating the term “gula aren” into “palm sugar” as seen from the source or material, the origin in making this sugar. The translation result should be “The sweet smell led us to a local home making *palm sugar*”.

1.2 Utilizing a very specific term for the generic term

Some terms contained in the source language are stated in general terms. These general terms must be equated with terms that are equivalent to the target language to avoid misinterpretation of the readers. However, if general terms in the source language are translated using more specific terms, it will cause errors in the translation results because the meaning of the source language does not match the target language. Therefore, readers identify different meanings or have distinctive understandings of the source language. This error is classified as confusion of sense relations in utilizing a very specific term for the generic term.

Data:

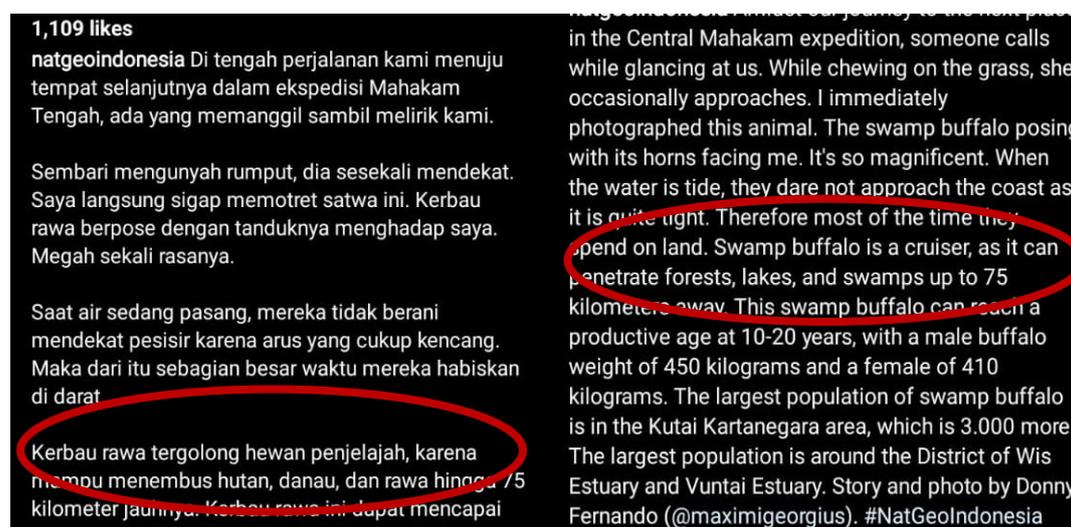


Figure 1.2 Natgeo Indonesia’s Post on 29th of September 2021

SL: Kerbau rawa tergolong hewan *penjelajah*, karena mampu menembus hutan, danau, dan rawa hingga 75-kilometer jauhnya.

TL: Swamp buffalo is a *cruiser*, as it can penetrate forests, lakes, and swamps up to 75 kilometers away.

(<https://www.instagram.com/natgeoindonesia/> on 29th of September 2021)

The translation above belongs to semantic errors in lexis and can be classified as the confusion of sense relations in choosing too specific term for a general term. The source language used a general term. Meanwhile, the target language used a more a specific term in translating the source language. According to KBBI online, the term “penjelajah” means an explorer of an unfamiliar area; an adventurer. While, based on Oxford Learner’s Dictionary 2021, “cruiser” refers to a large fast ship used in war; a boat with a motor and room for people to sleep, used for pleasure trips. The two tems

of “penjelajah” and “cruiser” have different meanings because the term “penjelajah” in the source language is actually has a general meaning because it is not focused on one object only. The term “cruiser” refers to a specific term as a type of warship used to explorer an area in war or nowadays, the term “cruiser” refers to a ship for pleasure trips. Therefore, the Instagram machine translation has failed to convey the general meaning of the term “penjelajah” which refers to the swamp buffalo. Instagram machine translation translated the term “penjelajah” specifically in the source language becomes “cruiser” in the target language. The term “cruiser” has a specific meaning because it is actually specialized for a ship. The Instagram machine translation is able to convey the meaning of source language correctly if translating the term “penjelajah” into “explorer” which is equivalent to the source language. The translation result should be “Swamp buffalo is an *explorer*, as it can penetrate forests, lakes, and swamps up to 75 kilometers away”.

1.3 Utilizing the inappropriate of two co-hyponyms

A hyponym is a term used to designate a particular member of a wider class. A hyponym refers to a word or phrase whose meaning is included in another more general word or phrase. These more general words or phrases are called hyperonyms or hypernyms. Words that are hyponyms of the same broader term are called co-hyponyms. If an error occurs in the selection of co-hyponyms in the translation, it will cause differences in understanding between the source language and the target language. This error can be classified as confusion of sense relation in utilizing the inappropriate of two co-hyponyms.

Data:

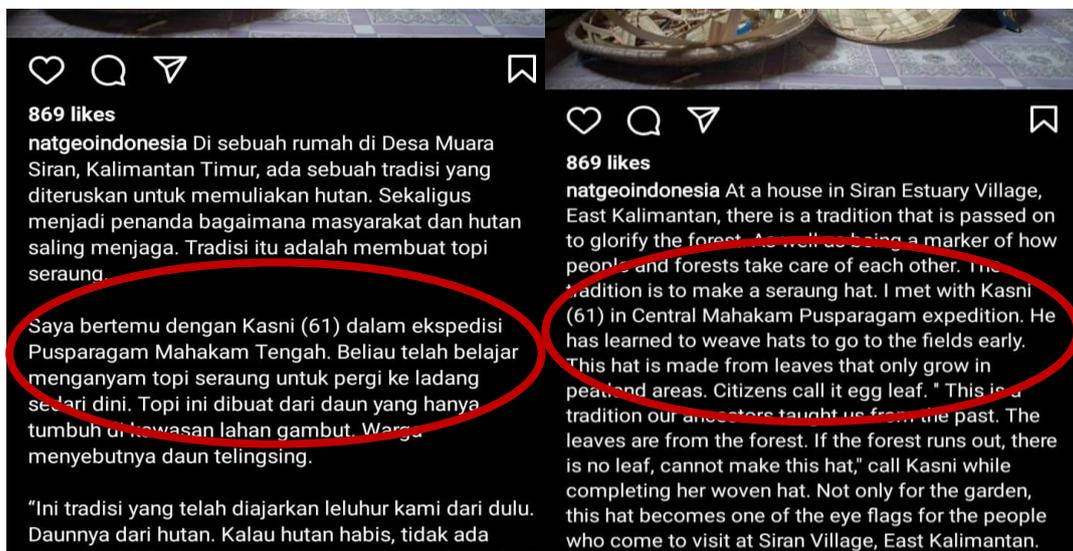


Figure 1.3 Natgeo Indonesia’s Post on 27th September 2021

SL: *Beliau* telah belajar menganyam topi seraung untuk pergi ke ladang sedari dini.

TL: *He* has learned to weave hats to go to the fields early.

(<https://www.instagram.com/natgeoindonesia/> on 27th of September 2021)

The translation result above can be categorized into semantic errors in lexis and classified as the confusion of sense relations in choosing inappropriate of two co-hyponyms. The fact, another hyponym which is more suitable in delivering the right meaning of the source language is available. The error in selecting appropriate hyponym from the source language to the target language leads to the alteration of the meaning and confusion in the context of sentence. According to KBBI online, the term “Beliau” refers to men or women. While, based on the context of the sentence, the term “Beliau” refers to a woman named Kasni. She is 61-years-old woman as a *seraung* hat weaver. However, according to Oxford Learner’s Dictionary 2021, the term “He” is a pronoun used for a man. These two terms refer to different meanings because based on the context of the sentence, Instagram machine translation has selected the inappropriate of two co-hyponyms. There are two hyponyms of the pronoun “Beliau” namely “She” and “He”. Instagram machine translation has translated the source language into target language without concerning to the context of the whole sentence. While, the use of pronoun “He” could not represent the term “Beliau” based on the translation above. It causes the error occurred in this translation. Instagram machine translation translated the term “Beliau” literally in the source language becomes “He” in the target language. The term “He” in the target language causes misinterpretation of the meaning to the readers in the source language. The Instagram machine translation could deliver the meaning of source language correctly if concerning to the context of the whole sentence. Therefore, the most appropriate pronoun to transfer the meaning of the term “beliau” based on the context of entirely on the sentence is “She”. The translation result should be “*She* has learned to weave hats to go to the fields early”.

1.4 Utilizing the inappropriate a pair of close-synonym

The words choice in translation is very important in producing accurate translation results. From the source language translated into the target language usually, it is necessary to search for an equivalent synonymous term in the target language. However, if the selected term is a synonym without having equivalent to the meaning contained in the source language, it will lead to misunderstandings and semantic errors. The error that occur in this translation are classified as confusion of sense relation in utilizing the inappropriate a pair of close-synonym.

Data:



Figure 1.4 Natgeo Indonesia's Post on 27th September 2021

SL: Sekaligus menjadi *penanda* bagaimana masyarakat dan hutan saling menjaga.

TL: As well as being a *marker* of how people and forests take care of each other.

(<https://www.instagram.com/natgeoindonesia/> on 27th of September 2021)

Semantic errors in lexis contained in the translation above and can be categorized as the confusion of sense relations in selecting an appropriate synonym. The source language was translated by an acceptable synonym into target language. According to KBBI online, the term “penanda” means an indication of something. Based on the context of the whole sentences the tradition that is passed on to glorify the forest is a sign of how people and forest take care of each other. However, according to Oxford Learner’s Dictionary 2021, “marker” refers to an object or a sign that shows the position of something or someone or something that marks. These two terms refer to different things because based on the context of the sentence, the term “penanda” in the source language is actually the term used to show that something exists. Consequently, the Instagram machine translation has failed in conveying the correct synonym from source language to the target language. Instagram machine translation has transferred the meaning from source language into target language without concerning to the context of the sentence. Instagram machine translation translated the term “penanda” literally into “marker” in the target language. Therefore, the appropriate synonym in translating the term “penanda” is utilizing the term “sign” which is equivalent to the source language. The translation result should be “As well as being a *sign* of how people and forests take care of each other”.

2. Collocational Errors

Collocation is two or more words in English that are tied together and thus creating a particular meaning. The combination of words in collocation usually forms a natural or commonly used phrase and generates a different meaning when separated from the words that compose it. The error that most often occurs in translating collocation is translating the constituent words in the source language literally so that it causes unnaturalness in the target language or tends to sound strange to native speakers of the target language. This error belongs to the confusion of sense relations in collocational errors.

Data:

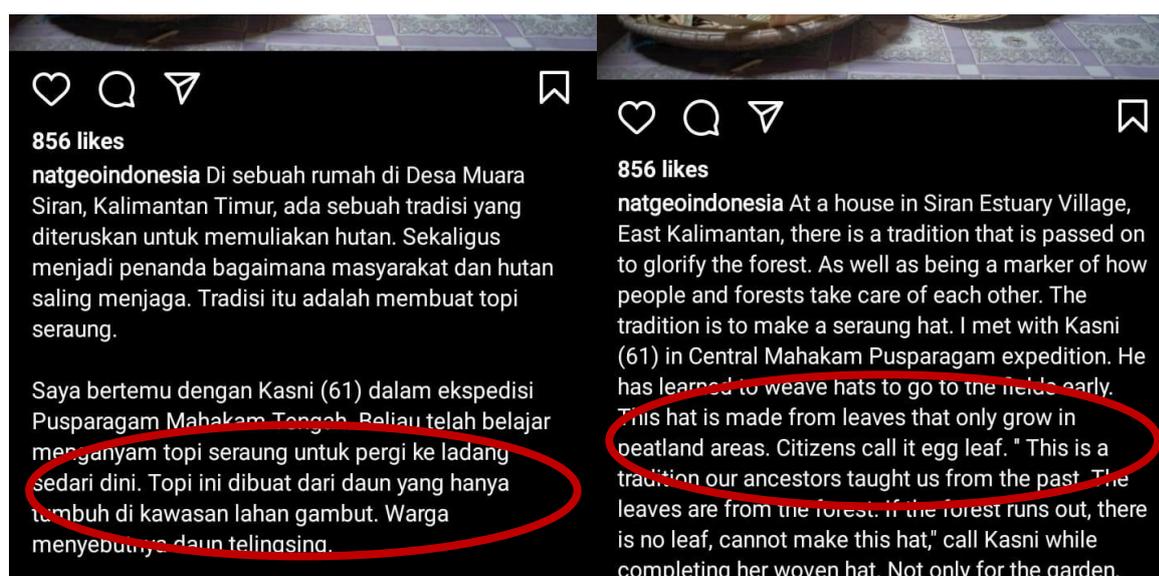


Figure 2.1 Natgeo Indonesia's Post on 27th of September 2021

SL: Topi ini *dibuat dari* daun yang hanya tumbuh di Kawasan lahan gambut.

TL: This hat is *made from* leaves that only grow in peatland areas.

(<https://www.instagram.com/natgeoindonesia/> on 27th of September 2021)

The target language above contains semantic errors in lexis and can be categorized as the collocational errors. The source language was translated by involving inadmissible collocation thus a mistake in the target language. The collocation of “dibuat dari” are closely affiliated with each other. According to KBBI online, the collocation of “dibuat dari” means a verb shows something has been produced using material. However, the Instagram machine translation translated the collocation “dibuat dari” into “made from”. The collocation of “made from” is a literal translation from the source language to the target language and sounds less common to native speakers of the target language due to the different forms of collocation in describing the collocation of “dibuat dari” in the source language. The collocation of “made from” could not deliver the meaning of “dibuat dari”. Instagram machine translation is able to convey the meaning of source language correctly translating the term “dibuat dari” into “made of”. According to Oxford Learner’s Dictionary 2021, the collocation composed of the verb “made” and the preposition “of” generates a new meaning namely the basic material in producing something or something has been created utilizing a substance or material, so

the first substance or item is totally changed. The translation result should be “This hat is *made of* leaves that only grow in peatland areas”.

CONCLUSION

Based on the findings in the previous chapter, it can be concluded that there are 238 data of semantic errors of lexis made by Instagram machine translation in the Translation of National Geographic Indonesia’s Instagram caption. There are two types of semantic errors in lexis. The first is confusion sense of sense relations can be divided into four sub-types; a) utilizing a generic term for the specific term, b) utilizing a very specific term for the generic term, c) utilizing the inappropriate of two-cohyponyms, and d) utilizing the inappropriate a pair of close-synonyms. The second type of semantic errors in lexis is collocational errors. From the 238 data, there are 170 (71.42%) data of confusion of sense relations; a) 19 data (7.98%) of utilizing a generic term for the specific term, b) 10 data (4.20%) of utilizing a very specific term for the generic term, c) 5 data (2.10%) of utilizing the inappropriate of two-cohyponyms, and d) 136 data (57.14%) of utilizing the inappropriate a pair of close-synonyms. There are 68 data (28.57%) of collocational errors. The most type that showed is the utilizing the inappropriate a pair of close-synonym with 136 data (55.04%). This phenomenon occurred because Instagram Machine Translation used an inappropriate translation procedure by choosing the lexis in the target language literally through a term that which have several synonyms which have different meaning in each use without considering the overall context of the caption. The least founded data is utilizing the inappropriate of two co-hyponyms with 5 data (2.10%). This phenomenon occurred because the data that consist of co-hyponyms are rarely found on Instagram machine translation and usually the data have been translated correctly.

REFERENCES

- Bahrami , A., & Salehi, M. (2018, October 22). An error analysis of journal papers written by Persian authors. *Cogent Arts & Humanities*, 5(1), 1-16. doi:10.1080/23311983.2018.1537948
- Bahasa, B. P. (2016). *KBBI Daring*. Retrieved from May 15, 2021, from <https://kbbi.kemdikbud.go.id/>
- Budiman, D., & Saifullah, A. R. (2020). The Translation of Relative Pronoun 'Yang' in 9 Summers 10 Automns, An Indonesian Novel: Syntactic and Semantic Studies. *Apollo Project*, 9(2), 25-37.
- Hatim, B., & Munday, J. (2004). *Translation: An Advanced Resource Book*. London and New York: Routledge.
- Hutchins, W. J., & Somers, L. H. (1992). *An Introduction to Machine Translation*. San Diego: Academic Press Inc.
- Instagram.com. (n.d.) *Instagram National Geographic Indonesia*. Retrieved from May 1, 2021, from <https://www.instagram.com/NationalGeographicIndonesia/?hl=id>
- James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. London and New York: Routledge.
- Katz, J. J. (1972). *Semantic Theory*. The University of Michigan: Harper & Row.

- Larson, M. L. (1998). *Meaning-Based Translation: A Guide to Cross-Language* . New York: University Press of America.
- Leech, G. (1981). *Semantics The Study of Meaning* . Middlesex: Penguin Books Ltd.
- Oxford Learner's Dictionaries*. 2021. Retrieved May 15, 2021, from https://www.oxfordlearnersdictionaries.com/definition/american_english/translation
- Shormani, M., & Al-Sobani, Y.A (2012). Semantic error committed by Yemeni university learners: classification and sources. *International Journal of English Linguistics*,2(6)
- TechTerms.com. 2014. *Instagram Definition*. Retrieved May 31, 2021, from <https://techterms.com/definition/instagram>