ERRORS MADE BY GOOGLE TRANSLATE AND ITS RECTIFICATION BY HUMAN TRANSLATORS

FARAH HANA AMANAH

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

ERRORS MADE BY GOOGLE TRANSLATE AND ITS RECTIFICATION BY HUMAN TRANSLATORS

FARAH HANA AMANAH

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF MASTER
OF ENGLISH AS A SECOND LANGUAGE

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **FARAH HANA AMANAH** [redacted]

Registration/Matric No: **TGB110016**

Name of Degree: **MASTER OF ENGLISH AS A SECOND LANGUAGE (MESL)**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

ERRORS MADE BY GOOGLE TRANSLATE AND ITS RECTIFICATION BY HUMAN TRANSLATORS

Field of Study: **TRANSLATION STUDIES**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signatur[redacted] Date: 13 February 2017

Subscribed and solemnly declared before,

[redacted]

Witness's Signature Date: 13 February 2017

Name: ASSOCIATE PROF. DR SU'AD AWAB

Designation: Supervisor

ABSTRACT

The use of bilingual versions of public communication such as notices letters, pamphlets or legal documents are the norm and often times are expected as Malaysia is a multiracial country. Having a bilingual version is said to cost a certain amount of money depending on the length of a text. This study aimed at investigating and evaluating the output of Google Translate which requires human editing and thus, to find out to what extent does Google Translate help in assisting human translation in particular with Bahasa Melayu (BM) into English translation. In this study, four texts namely notices and pamphlets were used. Three female translators who edited the output of Google Translate were tasked to check the translations. A hybrid taxonomy on error analysis proposed by Keshavarz (1999) and Vilar et al. (2006) in terms of lexical, grammatical (syntactic) and semantic errors is used to identify the errors in Google Translate output. The findings indicated that most of the errors were in the grammatical category such as missing prepositions, missing possessive markers, omission of pronouns and/or incorrect use of lexical words. It is also revealed that using a different choice of word will affect the meaning as well. Hence, the three human translators were needed to rectify the errors, although the three of them had different approaches in correcting the errors. Despite that, it can be said that Google Translate is able to produce a good translation output when the text uses simple sentence structures and thus, Google Translate can aid or assist students or language practitioners in getting a general idea of the text but human evaluation is still required especially in keeping true to the meaning of the original text.

ABSTRAK

Penggunaan versi dwibahasa di Malaysia untuk tujuan komunikasi seperti surat, risalah atau dokumen undang-undang adalah penting dan sering kali mendapat perhatian kerana Malaysia terdiri daripada berbilang kaum. Terjemahan ke satu bahasa lain seringkali dilakukan dan ia memerlukan sejumlah kos yang tertentu bergantung kepada panjang teks. Oleh itu, kajian ini bertujuan untuk mengkaji dan menilai hasil terjemahan *Google Translate* yang memerlukan penilaian manusia, selain untuk mengkaji sejauh mana *Google Translate* dapat membantu dalam kerja penterjemahan khususnya terjemahan bahasa Melayu (BM) ke dalam bahasa Inggeris. Empat teks iaitu notis dan risalah telah digunakan dan tiga penterjemah perempuan yang berpengalaman dilantik untuk menyunting hasil terjemahan *Google Translate* berdasarkan kesalahan-kesalahan yang dilakukan oleh *Google Translate*. Kajian ini telah menggunakan pendekatan *'Error Analysis'* yang dicadangkan oleh Keshavarz (1999) dan Vilar et al. (2006) yang menganalisis dari segi leksikal, kesilapan tatabahasa (sintaksis) dan makna. Dapatan kajian menunjukkan bahawa kebanyakan kesilapan adalah dari kategori tatabahasa seperti kata hubung atau kata tempat, hilang penanda posesif, peninggalan kata ganti dan / atau penggunaan kata yang tidak betul dari leksikalnya. Selain itu, hasil kajian mendapati penggunaan pilihan perkataan yang berbeza juga menjejaskan makna. Oleh itu, ketiga-tiga penterjemah diperlukan untuk membetulkan kesilapan yang dilakukan oleh *Google Translate* dan mereka menggunakan pendekatan yang berbeza dalam membetulkan kesilapan itu. Namun, ia boleh disimpulkan bahawa terjemahan dari *Google Translate* masih mampu menghasilkan terjemahan yang memuaskan apabila teks menggunakan struktur yang mudah dan ringkas dan dengan itu, *Google Translate* dikatakan boleh membantu pelajar atau pengamal bahasa untuk mendapatkan gambaran umum teks walaupun beberapa makna masih memerlukan penilaian manusia.

ACKNOWLEDGMENT

First and foremost, I would like to express my gratitude to Allah S.W.T, the Al-Mighty for His blessings giving me the opportunity and ability to complete my dissertation writing. After a few years through ups and downs that I faced, I have finally completed my study for a Master's Degree. Indeed, I have learned the true meaning of patience and perseverance throughout my journey as a postgraduate student that makes me become a better person.

I would also like to express my million thanks to the one that I indebted the most, my respected supervisor, Associate Professor Dr. Su'ad Awab. The continuous guidance, moral support, dedication and advice as well as her valuable knowledge that she gave me ever since I started to become her supervisee are invaluable. I can only keep her in my prayers. Truly, I am fortunate for being her supervisee as she has always been consistent in asking my progress and gave constant feedback on my dissertation. The time that I had with her will always be cherished and remembered.

Additionally, I would like to express my heartfelt appreciation to my family especially to my parents; Amanah Neemat and Anitah Haji Salim. Both of them are the pillars in my life that have always believed in me and trust me that I can face all the obstacles while completing my Master's Degree. Their constant prayers, moral and financial support, motivation as well as advice that they have given to me motivated me to complete my study. It has always been their dream to see me succeed and seeing them happy is the true definition of an accomplishment as they are the reason of what I am now. Thank you Ibu and Abah for guiding me and for always being with me during my bad and good days throughout my study.

To my beloved husband, Mohd Syarizan Hamzah who has never failed to support me in every physical and mental aspects. The man who always inspires me to keep and continue doing what I have started years back until I completed my dissertation. Thank you Abang, for the endless and tremendous support that you have shown to me. I pray to Allah S.W.T, the Al-Mighty God to answers our prayers for having a beautiful family soon.

I would also like to thank to my dearest siblings, friends, supportive colleagues, aunts and uncles who have always keep me in their prayers and keep encouraging me to never give up. Not forgotten also to those who have helped me directly or indirectly in completing this dissertation.

Certainly, each one of them means a lot to me. They have witnessed my struggles and obstacles in completing this dissertation throughout the years. I could never have to thank them enough. I can only pray to Allah S.W.T to reward them for what they have done. Finally, this dissertation is dedicated to my beloved parents and husband.

TABLE OF CONTENT

PREFACE

CHAPTER 2 LITERATURE REVIEW

CHAPTER 3 RESEARCH METHODOLOGY

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

UT	Unit of Translation
HT	Human Translator
MT	Machine Translation
ST	Source Text
TT	Target Text

CHAPTER 1

INTRODUCTION

1.0 Background of the Study

Malaysia is a multiracial country with different cultures and languages. There are different languages and dialects such as Bahasa Melayu, English, Mandarin, Tamil and others which are used to communicate either in written form or spoken discourse. In Malaysia, bilingual versions of public communication such as notices, letters, pamphlets and legal documents are the norm and often times are expected. It can be seen that any information in Bahasa Melayu (BM) often has its equivalence in English. This is to ensure that everyone will get the information that is being conveyed. Hence, there is a need for a quick resource to produce the translation since there are high demands for translating the information from BM to English. Translation can be defined as the act of transferring the meaning or a unit of language from a source text to the target text (Newmark, 1988, p. 12). In other words, translation is transmitting the source text (BM) into the target text (English) by individuals who have a good knowledge in both languages.

In order to translate a piece of work from BM into English, there are numerous web-based machine translations (MTs) which are readily available such as Systran[1], Microsoft Bing Translator[2] and Google Translate[3] (Jia, Junmoa, Dongqing, Tiejun & Shen, 2012). The use of one of the web-based MTs helps to reduce time and is less costly compared to hiring a professional translator. According to Nurul Hidayah Abdillah (2014) who works for Institut Terjemahan & Buku Malaysia (ITBM), the cost of translation depends on the length of sentences and the types of text, where according to

[1] http://www.systranet.com/translate
[2] http://www.bing.com/translator/
[3] http://translate .google.com.my/

her, the usual cost would be twenty to thirty cents per-word[4]. Since the cost of translation can be pricy, one of the best alternatives is to employ a web-based MT engine. She also commented that some translators use Google Translate in order to get a preliminary gist of a source text (ST) especially a complex text. Similarly, Aiken and Balan (2001) state that the use of machine translations is to get a quick translation of a foreign language in order to help the translators get the best meaning in the target language. The translators' work become easier as they simply have to edit the translated text in order to ensure that the intended meaning in the source text is preserved[5].

The advancement of technology nowadays, provides a room for the translators or anyone to practice translating a short or simple sentences and phrases by using a web-based machine translation. It is quick, shortens the time to produce the output of target text and less costly, however, the accuracy of the output produced by such machine translations is debatable and questionable.

1.1 Statement of Problem

Translating is a taxing task, which needs a translator to spend his or her time on a work that he or she translates. The translated works need to be done in a given time. Due to time constraints, the use of immediate and automatic web-based machine translation is a solution apart from other translation aids such as dictionary and translation software. It is helpful to have a translated version of the target text produced by using one of the machine translations which needs less editing depending on the text type. Handschuh (2013) conducted a study on German-English translation using four different online machine translations (ie: Google Translate, SYSTRAN, Bing and Babylon) and found out

[4] Personal communication with Nurul Hidayah on 20th April 2014
[5] Personal communication with Nurul Hidayah on 20th April 2014

that there are still errors found in the target text output produced by the machine translations. The original meaning is often and not retained when longer texts are translated.

Some of the web-based machine translation issues are: How far Google Translate, a popular machine translation, can translate accurately; what types of errors made by Google Translate which need to be edited by human translators or at what level can Google Translate translate better. The translation produced by Google Translate seems to be awkward sometimes, hence, there is still a need to have a human touch to ensure the quality of output of the target text. Hutchins and Somers (as cited in Talal Muhammad (2012) discuss a few questions on how well this machine translation (as in Google Translate) produced its raw output, what errors need to be improved or how effective or practical is its approach. Likewise, Juan (1994) suggests that the problem of using machine translation lays on the output of language, meaning as well as the cultural differences that are produced by machine translation, as some of the languages might not have culture similarities like the use of figurative language.

This study is a descriptive analysis of Google Translate output which requires human editing and thus, to find out to what extent does Google Translate help in assisting human translation in particular with BM to English translation. Keshavarz's and Vilar et al.'s taxonomy of error analysis will be used in analyzing the translation outputs. The translation outputs will be discussed on three (3) aspects that are in terms of lexical, grammatical (syntactic) and semantic. The lexical aspect is related to word choices while grammatical (syntactic) focuses on the structure of a sentence and semantic aspect on the other hand, looks at the meaning of a text. This taxonomy or framework will be further explained in specific in Chapter 2 and Chapter 3. Thus, this study aims to fill the gap of research on the use of Google Translate in translating different types of text.

1.2 Significance of the Study

This study is significant as it will provide useful information on how web-based machine translation can be used effectively when translating different types of texts such as notices, advertorials and pamphlets from BM into English. The findings from this study could clarify on how Google Translate works and its reliability in producing output from the source text at word level, sentence level or in terms of its syntactic, semantic and lexical patterns or features.

Next, the results gathered from this study can be used as a future reference for training translators to adopt Google Translate in translating which would help them to get a quick result of translation even though they still have to do some editing of the output. In order to improve the quality and accuracy of machine translations, this study is beneficial for experts in machine translations. The researcher believes that the experts will improve the accuracy and reliability of machine translations by adding more words or vocabularies in the machine translations' bank or re-structuring the structure of machine translations of different languages.

Thus, the findings from this study are expected and hoped to give guidelines for different fields of language learning or expertise as well as those who are interested in translation.

1.3 Objectives of the Study

This study aims at investigating the errors produced by Google Translate when translating different texts (in which this study focuses at informative and vocative texts) as well as examining the produced output, which is then rectified by human translators in order to improve the quality of the translation.

1.4 Research Questions

In order to achieve the objectives, the following research questions were formulated:

i) What are the typical errors produced by Google Translate when translating informative and vocative texts from BM into English?

ii) How are the errors rectified by the human translators to improve the quality of the translation?

1.5 Scope and Limitations of the Study

There are several limitations of this study on how Google Translate can aid human translators in producing output from BM into English. Firstly, this study used different texts that are notices and pamphlets which are customarily issued in both BM and English. Initially, the researcher identified only one text that is notices. However, due to the restriction in analysing the data, the researcher added another text that is pamphlets. Therefore, the results gathered from this study are limited to the text types that have been mentioned earlier and cannot be compared to other text types as in technical texts or figurative texts.

Secondly, there are only three female participants involved in this study. These participants are free-lance translators who have approximately three to four years of experience in translation. All of them were chosen based on the knowledge of both languages (BM and English) as well as their experience in translation. The number of participants is small as the editing and corrections done by them are varied and thus required careful scrutiny. The results gathered cannot be generalised to different participants or learners.

Finally, this study only adopts one of the web-based machine translations which is readily available that is Google Translate. Google Translate is well known and familiar to most computer users, thus it is chosen for this study. Google Translate consists of and can translate 80 languages including bahasa Melayu and English. The target text output (English) produced by Google Translate might differ with other machine translations.

1.6 Conclusion

This chapter has discussed the overview of the study by providing the research problem, executing the research questions, stating the significance of the study as well as identifying the scope and limitations of the study. There will be four more chapters for further discussion throughout of this dissertation writing which include *Literature Review* in Chapter 2, *Methodology* in Chapter 3, *Findings and Discussion*s in Chapter 4 and the final Chapter of 5 will be on the *Conclusion and Recommendation.*

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

In this chapter, relevant literature pertaining to this study will be further expounded and discussed. It includes previous and similar studies conducted by other researchers on this study. This chapter, therefore, presents the related theories which will expand the context and background that are related prior to this study.

2.1 Translation

A layman may define translation as 'translating from one language to another language'. For experts, particularly those who are involved in translation, have proposed several definitions of translation. Newmark (1988) asserts that translation is rendering the meaning of a source text into another target text in a way that the author or translator intended the text to be. Translation is also defined based on three bases of knowledge which are general subject field, the product and the process (Munday, 2008, p. 5). Meanwhile, Jakobson (1959) as cited in Venuti (2000) categorises translation into intralingual (rewording), interlingual (proper translation) and intersemiotic. In other words, translation deals with the process of translating that involves the changing of an original text to the product of the source text.

Nida & Taber (1964) agree that translating is producing the receptor language or text that has equivalent meaning as the source language or text. Nida and Taber (1969) further elaborate that the target text is transferred or transformed into the surface structure of the target text that involves three stages (analysis, transfer, and re-structure).

Furthermore, Catford (1965) proposes that translation is the substituting process of a text (source text) into another text (target text). Moreover, Larson (1998) defines translation as the act of translating the source text into the target text in a natural form in meaning. Bell (1991) adds a new dimension of translation by looking at it as a process of transforming the source text meaning into the target text that is related to psycholinguistic and sociolinguistic factors.

It can be concluded that translation involves two languages where the former is from the source text while the latter is the output of the target text which is the second language of the text. Therefore, to obtain a translation of quality, the output has to produce a similar work or renders a similar meaning to the one in the source text.

2.2 Brief History of Machine Translation (MT) and Computer-assisted Translation (CAT)

Machine Translation (MT) is a scientific discipline which involves the application of computer programmes to translate different texts or languages into another (Ping, 2009). It was developed in the 1950's as a computerised system that performs automatic translation (Hutchins, 1995; Balkan, 1992). In the beginning, the system works when the whole texts in a source language (SL) are translated into a target language (TL) as a single task without human intervention. The source text output produced by machine translation is known as 'raw output' as it provides a quick translation of the original. These raw outputs usually offer informative translation. It means, the output produced by the machine translation only provides surface translation of the target text without human involvement, as it is a statistical machine translation. It is deemed to translate what is inserted in the system. Machine translation is one of the oldest applications and has been used in computer science. Nonetheless, due to

globalisation and needs for transferring knowledge and information, it has been used in language and linguistic fields as well (Talal Mohamad, 2012).

Additionally, the main objective of machine translation during the early stage was to replace human translators as it can do the translation work. However, unsatisfactory output produced by the system and problems that could not be solved due to lexical ambiguities produced by machine translation made the enthusiasm among the experts decline (Juan, 1994). Many companies that developed machine translation at the early age started to admit that the systems were no longer able to produce perfect translation. Therefore, due to the failure system and unsatisfactory output produced, it has led to the development of computer-assisted translation (CAT).

Subsequently, computer-assisted translation enables human intervention in the machinery system, thus, helping translators to work quickly (Craciunescu, Sales & O'Keefe, 2004). It helps to assist human translation's work and gives a human translator an extended control over the process.

2.2.1 Web-based Machine Translation (MT): Google Translate

In the era of globalisation and nearing Vision 2020, the use of internet has been rapidly increasing. The machine translator's vendor has started to offer online translation services which are available readily and allowing everyone to have access to the systems. Google Translate works on the principle of statistical machine system which allows it to guess and decide on the most appropriate translation output (Retrieved from http://translate.google.com.my/about/int/en_ALL/). It relies on a vast body of stored translated work database done by human translators. The process starts when a text is submitted online, Google Translate will go through this database and produce what it deems as the best approximation of the target text. In other words,

'Google Translate' acts as a new learner of a foreign language which stores new vocabularies and grammatical rules.

Besides that, different languages have different grammatical forms. For some languages, Google Translate has fewer translated documents and, thus, the quality of the translations will vary from other languages and language pairs. An administrative from Google commented that;

> *" our translations are not always perfect, but by constantly providing new translated texts we can make our computers smarter and our translations better"*
>
> (Retrieved from http://googlesystem.blogspot.com/2010/08/)

Currently, Google Translate is able to translate over 80 languages including the Malay language. There are various web-based machine translations systems available online such as Systran[6], Baidu[7], Microsoft Bing Translator[8] and Google Translator[9]. Nonetheless, Google Translator is the most popular machine translation viewed by 200 million users on average (Jia et al., 2012). Hence, it is relevant to use Google Translate as one of the aids that is preferred by users in translating the source text to the intended target text. In fact, the use of Google Translate in few studies has shown that Google Translate has gained much interest among users (Karami, 2014; Aziz, Sousa, & Specia, 2012; Komeili, Farughi, & Rahimi, 2011).

[6] http://www.systranet.com/translate
[7] www.baidu.com
[8] http://bing.com/translator/
[9] http://translate.google.com.my/

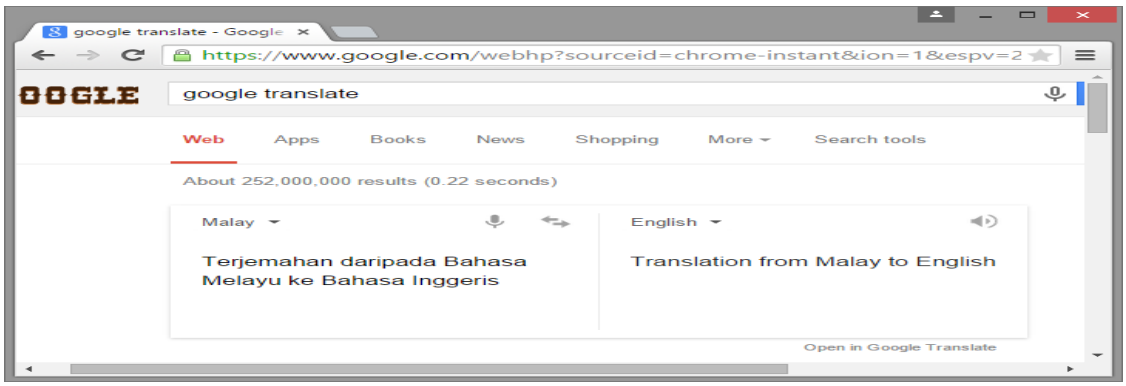

Figure 2.1: Features of Google Translate when translating ST to TT

Figure 2.1 above shows the feature of Google Translate when translating a source text into a target text. As mentioned above, a text is inserted into the system (Google Translate) and after a few seconds, the translated version of the target text is produced. It provides a raw output of the source text and where necessary, changes are deemed to be made by human translators in order to ensure the meaning of the text is relevant to the readers.

2.2.2 Meaning and Machine Translation

Abdel Monem et al. (2009) state that machine translation is used in multilingual translation with an interlingual approach. It involves two steps. The first step is the meaning of the source text is represented in an immediate language-independent (interlingua) and from those meaning representation, the sentences of the target text are formulated or generated. This is due to the reason that language complexity at both the morphological or lexical level and syntactic level that affects the semantic meaning of the text. For instance, the use of prefix or suffix in the Arabic language is semantically related to the root word (base form) and it will provide different meaning if translated wrongly.

Wilks's (1972, p. 3) views the meaning and machine translations look at the meaning of a structure of a language. This is where the grammatical structure of the input and output of languages is rearranged and then the input words are translated on a one-to-one basis of the output (target language). Thus, he points out that each Russian word in his study is based on one-to-one translation of an English word.

Toma (1959, p. 249) says that machine translation systems need to have a similar idea to human translation when translating a source text into a target text. It is when its source text is read to identify what individual words mean, the role for each word plays in any particular sentence as well as the denotation of any word is affected by its context. He further explains that the machine translation needs to have adequate vocabularies, syntactic analysis programs, semantic analysis procedures as well as the target language synthesis analysis. Additionally, all languages that are translated by machine translation have common and similar properties. The common properties of languages include words to express an action or equivalence to it, a relationship between the action and that which performs the action and a relationship between the action which is affected by the action. Therefore, the whole translation procedure includes lexical, syntactic and semantic properties when a text is read into the computer for translation processing. Some of the information is supplied by the machine translation dictionaries and other information is supplied from analysis and synthesis of the text. If some of the words are not found in the dictionaries, the output of the target text may be omitted or rendered the same word as in the source text. For example, in unit translation 26 (Text 2), *"Untuk memastikan* ***sekuriti*** *dan keselamatan premis (berkenaan imej-imej CCTV)".* The dictionary sources in machine translation is unable to provide equivalence word for 'sekuriti' and therefore, maintains the same word in the source text.

Delavenay, 1960 (p. 51) states that the meaning of the outputs produced by machine translation are imposed by choices and meaning of the words, the forms that words may take and the relationships that exist between them (i.e semantic, morphology and syntax). In order to identify the meaning of a source text, the machine must identify each of the grammatical forms and syntactic structures of a target text, making it is a comprehensible sentence in that language. She provides an example from an English text which is translated into Russian. Both of the texts are analysed and synthesised in terms of its vocabulary, parts of speech, syntax as well as its meaning. She asserts that English words can have more than one meaning and thus sometimes makes it difficult to choose from among the several meanings provided by the machine translation to the correct Russian word which is equivalent for the given context. She even explains that incorrect placement of syntax also results to wrong translation as it does give impact to the meaning of the target text.

To conclude, the meaning of translation outputs which are produced by machine translations are connected with the three properties mentioned above: lexical, syntactical and semantic. Some of the errors that are produced by the machine translation affect the meaning due to wrong word order or incorrect use of word choices depend on its functions and text types.

2.3 Text Types and Text Functions: Pamphlets and Notices

There are different kinds of text that can be translated. These texts serve different functions and purposes. According to Reiss (1971 / 2000, p. 160), there are three different functions of text typology for communicative functions that are used which correspond to the type of texts: expressive, informative and operative functions

or communication (refer to Figure 2.2). Expressive communication is a creative composition where the author uses an aesthetic or figurative language to express the message. Meanwhile, informative communication or plain communication transmits information, knowledge and opinions that are logical and the content of the information is the sole focus. Operative communication, on the other hand, aims to appeal or persuade the readers to act in a certain way.

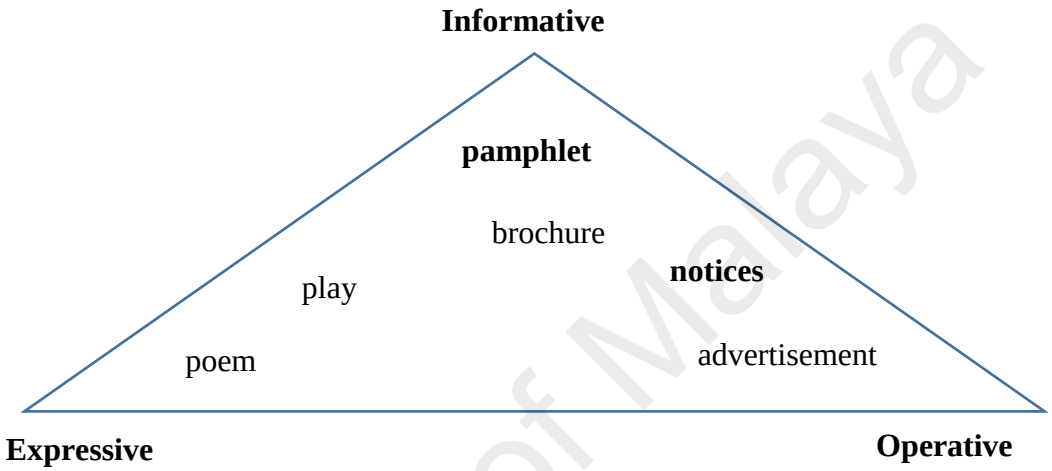

Figure 2.2: Reiss' s text types (Adapted from Chesterman, 1989)

Based on Reiss's text typology, notices and pamphlets that are used for this study are identified as an operative text since its main function and purpose are to provoke and encourage behavioural responses from the target readers by persuasion (reader-and-effect oriented). Additionally, all of them also convey essential information (informative) and, thus, content-oriented.

Similarly, Newmark (1988) adopts a functional theory of language which believes that there are three major purposes in using the language. The three purposes are expressive, informative and vocative as shown in Table 2.1. According to Newmark (1988), the expressive function is the utterance of the writer or speaker. He believes that the writer or the speaker uses the utterance to express his or her feelings towards any response such as in poem, lyrics, novels or plays. Meanwhile, the informative function

in a text refers to the facts of a topic, which is reality outside the community. Informative function also includes ideas or theories. It can be seen that informative text usually has a standardised format such as textbooks, minutes, and articles in newspaper or scientific papers. The last function in using the language as said by Newmark (1988) is the vocative function. The 'vocative' term is used to 'call upon' the readership to act, think or reflect in the way that the text is presented.

Table 2.1: Language functions, text-categories and text-types (Adapted from Newmark 1988)

Function	**Expressive**	**Informative**	**Vocative**
Core	Writer	'Truth'	Readership
Author's status	'Sacred'	'Anonymous'	'Anonymous'
Type	• Literature • Autobiography	• Scientific • Commercial • Textbook • Article • Minutes • Pamphlets	• Notices • Instructions • Propaganda

As been explained by Reiss (1971) and Newmark (1988), the main purposes in using the language are the expressive, informative and vocative or operative functions. Several texts can act and emphasis on two functions as in informative and operative. Nevertheless, expressive function has no place in the informative or vocative text (Newmark, 1988). Therefore, pamphlets and notices used in this study serve two functions which are informative and operative (vocative).

2.4 Translating Text: Unit of Translation (UT)

It is a fundamental concept of unit of translation that is always being debated among the experts. Different theorists have defined many definitions on unit of translation. One of the definitions given by Shuttleworth and Cowie (1997) on unit of translation is "a term used to refer to the linguistic level at which source text is re-codified in the target text". According to Barkhudarov (as discussed in Shuttleworth and Cowie, 1997), unit of translation is "the smallest unit of source text or language that has an equivalent in the target text or language". He further explains that morphemes, words, phrases, sentences or entire texts can be considered as a unit of translation and the most appropriate unit of translation for him is the wording of the source text.

However, the argument lies in the length of the unit of translation and between literal and free translation. Literal translation focuses more on individual words (morphemes). Thus, unit of translation is considered as short as words. On the other hand, free translation aims on a longer sentence (Hatim and Munday, 2004, p. 231). It depends on a translator to work on a smaller unit or work on a larger unit in order to convey the meaning of the source text.

Vinay and Darbelnet (1958) in Venuti (2000) introduce the most clear-cut approach relating to the unit of translation. They introduce the concept of word as a basis for unit of translation and identify unit of translation as a combination of 'lexicological unit' and 'unit of thought'. Lexicological units contain a lexical element which is grouped together and form a single element of thought, for instance; *'to give the people'*. On the other hand, unit of thought is when a translator has to consider the whole structure and context of the source text. Vinay and Darbelnet (1958) consider that the smallest unit of utterance where the signs are linked as in a word and concept should not be translated individually or separately. The sign is arbitrary in nature and derives

meaning with other sign in the same language. For example, ‘a tree’ in English signifies plant with a trunk and also can be signified as a bush. In addition, the relationship between units of translation and words inside the text by Vinay and Darbelnet (1958) can be categorised into (1) simple units, (2) diluted units and (3) fractional units (Esteki, 2010). A simple unit of translation is the simplest unit where it has an equal number of words in both the source text and the target text. For diluted unit, it contains several words that generate a lexical unit. Fractional unit is a fraction of words which is broken down into a smaller part in which this type of unit translation as proposed by Vinay and Darbelnet (1958) is consisted of. In contrast, Newmark (1988) focuses on the “sentence as a natural unit of translation”. He proposes that paragraphs and texts are the higher units of translation while sentences, groups, clauses and words as the lower units of translation.

Thus, the use of translation units as suggested by Vinay and Darbelnet (1958) is used to produce the outcome of the texts when translating from BM into English. The texts used for this study consist of sentences as well as phrases. This helps the researcher to analyse the output according to the unit of translation as it was numbered (UT1, UT2, UT3 etc.). The source text output was first translated as a whole text by Google Translate in order to ensure that Google Translate was able to provide the translated version as in one text, as mentioned before, the text consists of some phrases. Once translated, the target text output was numbered by using unit of translation. The rationale of using unit of translation (UT) is to identify the errors made by Google Translate when translating the pamphlets and notices (informative and vocative) sentence by sentence and not as a whole chunk of the text. This is because the machine translation system usually processes the largest unit of the text which is the sentence (Koponen, 2010) apart for some clauses or phrases, which are in bullet points.

2.5 Error Analysis in Translation

The use of error analysis (EA) is usually found in second language learning (Presada & Badea, 2014; Mogahed, 2011; Matsuzaki et al., 2015). Brown (2007, p. 257) distinguishes 'mistakes' and 'errors' where the former refers to a performance of errors made by the speaker or learner in a spontaneous situation which can be self-corrected. On the other hand, 'errors' refer to the competence of the learner where it reflects the understanding of the learner of a language and it cannot be self-corrected. To illustrate this, an English learner says, "*John cans sing*" in few occasions and further examination of the learner reveals that the learner is unable to distinguish modal verb from other verbs. Brown (2007) continues that although errors cannot be self-corrected, they can be observed, analysed and classified and this process is called error analysis. This theory can be applied in finding and observing the errors made by machine-aided translation which is later on edited by human translators. Similarly, Corder (1981, p. 45) defines error analysis as the main process in gaining second language learning where the learners faced or made errors or mistakes.

Over the years, the use of error analysis has gained much attention in translation as it was studied by many researchers (Vilar et al., 2006; Eftkhar and Nourey, 2013; Hadis and Mahmood, 2016; Koponen, 2010). The theory of errors analysis is not only applied to the learners of language learning, but also to errors that are found in machine translation. Error analysis is the identification and classification of an individual error in the use of machine translation (Stymne & Ahrenberg, 2012, p. 1785). It helps to evaluate the output of the target text produced by the machine-aided translation. Keshavarz (1999) as cited in Hadis and Mahmood (2016) explains that error analysis refers to the process of collecting the samples, identifying the errors and classifying them into which category they belong to, and finally evaluating the errors.

Gass and Selinker (1994, p. 57) also identify six steps on how to analyse errors based on the error analysis model. The six steps are to collect the data, identify the errors, classify the errors, quantify the errors, analyse the sources of errors and the final step is re-mediating the errors. Brown (1994, p. 128) holds the same idea about error analysis framework where he agrees that the first step involves selecting the target text output which is then followed by identifying the errors made. Later, the errors are classified and explained.

With the use of the error analysis theory in translation, it allows the researcher in this study to identify and classify the errors when examining the target text output produced by Google Translate.

2.5.1 Translation Errors Category

To evaluate the errors produced by Google Translate, there is a need to classify and categorise the errors. Vilar et al. (2006) proposed that there are five (5) categories of errors. There are inflectional errors, word order error, missing words, extra words and incorrect lexical choice. Inflectional errors refer to errors that occur due to the incomplete form of the correctly generated word. Word order errors refer to a word which occurs both in the reference of source and target text languages. It refers to syntactic ordering where it distinguishes between local or long range re-orderings. The distinguishment is made to express the difference between re-ordering words only in a local context which is within the same syntactic chunk or not. For instance, the English-Chinese language pair and vice versa. Missing words on the other hand, are words that are deleted or omitted in the target text output. Extra words are a word which occurs as insertion in the target text which is unknown to the source text and it does not

share any reference to the source text. The last category or error suggests by Vilar et al. is lexical choice. It is a word that belongs neither to inflectional errors, missing nor to extra words. This is when the word is translated wrongly which affects the meaning.

Bojar (2011) has a similar classification of errors which is adapted from Vilar et al.'s (2006) framework. He divides the errors into four (4) categories which are bad punctuations, missing words, word orders as well as incorrect words.

Presada and Badea (2014) adapt Dagneaux et al.'s (1996) taxonomy of error identification. They lay down thirteen (13) errors including subject omission, verb tense, errors related to continuous aspect, sequence of tenses, conditional clauses, word order, uncountable nouns, prepositions, double negation, indefinite articles, word omissions and lexical confusion.

Elliot et al. (2004) conduct a study on French-English translation using four different machine translations where the outputs later are edited by a post-editor. The errors are segregated into an inappropriate, untranslated, incorrect, unnecessary as well as omitted words.

Koponen (2010) categorises the error classification into five (5) classifications. These include omission (where the source text is not conveyed by the target text), addition (the target text word is not present in the source text), untranslated concept (this is when the source text word cannot be translated and it appears in the target text output). Further explanation on the classifications is mistranslated concept (where the target text has mistranslated the source text in which the meaning has been distorted) as well as substituted concept (where there is equivalence of lexical choices from the source text).

Another classification of errors as suggested by American Translators Association (ATA) (2010) involves twenty-two (22) types of errors for classifying and grading. These errors include incomplete passage, illegible hand writings, and misunderstanding of the source text, mistranslation of the target text output, addition or omission. Besides that, they also count errors in terminology which include word choice, register, free translation, literal or word-to-word translation, false cognate of the source text form or incorrect use of word choice. Additionally, inconsistency in translation, ambiguity, grammar, syntax, punctuation, spelling, accents and other diacritical marks, the use of upper and lower case, word form, usage as well as style are classified into ATA's errors. These errors were useful and used by the translators as a guideline when translating different text types that serve different functions.

Keshavarz (1999, p. 11) believes that the errors are from lexico-semantic errors, wrong use of tenses, errors in the use of verb groups, errors in the word orders, incorrect use of prepositions, incorrect use of active and passive voice as well as errors related to the use of articles. He further explains that the errors can be specified into three sub-categorisations which are grammatical, semantic and pragmatic errors.

Hossein and Saleh (2013) categorise errors into four (4) broader categories, namely, syntactic, semantic, pragmatic and translation-specific errors. These errors taxonomy proposed by them is adapted from ATA (2010) and Keshavarz's (1999) linguistic taxonomy of errors. They further classify each of the categories with specific errors which is listed as follows:

i. Syntactic error: grammar, syntax, punctuation and usage

ii. Semantic error: addition or omission, terminology, free translation, word-to word translation, false cognate of the original form, ambiguity, accents, upper and lower case, word form, spelling

iii. Pragmatic error: misunderstanding of the source text, mistranslation of the target text output, register and style

iv. Translation-specific error: incomplete passage, inconsistency

Based on the errors category that are mentioned above, this study adopts both Keshavarz's (1999) and Vilar et al.'s (2006) model of errors identification taxonomy which has been used in other studies too (Hadis and Mahmood, 2016; Eftekhar and Nouraey, 2013; Hossein and Saleh, 2013; Angela Costa et al., 2015; Popoviv and Ney, 2011; Koponen, 2010). Similarly to Hossein and Saleh (2013), the researcher classifies the errors into three (3) categories, namely, lexical, grammatical and semantic/ pragmatic errors. Each of the categories is later specified with a detailed error. Table 2.2 shows a proposed model which has been adapted from Keshavarz's (1999) and Vilar et al.'s (2006) framework of error categorisation.

Table 2.2: A Hybrid Model of Error Classification

1.	Lexical Errors	- wrong word choices - different word choices
2.	Grammatical (Syntactic) Errors	- tenses - prepositions (missing / wrong preposition) - articles (missing / wrong article) - word reference (missing / wrong pronoun , missing possessive marker / missing relative pronoun) - wrong word form - wrong word order (SVA, gerund, conjunction)
3.	Semantic Errors	- errors in meaning - collocations

2.5.1.1 Lexical Errors

It consists of a list of words that belong to a syntactic category such as noun, verb, adjective and other parts of speech (Koemili et al., 2011). It contains errors that are related to wrong word choices in which to some degree will affect the meaning of the text. When the output of the target text is unable to identify the equivalent word as in the source text, the word is then translated using the wrong lexical item. For example, *"Perubahan dalam* ***'corak'*** *pembuangan air kecil yang berterusan dalam jangkamasa 2-3 bulan"* where *'corak'* is translated as *'twist'* (Example from Text 1). Thus, it can be observed that lexical interference can distort the meaning of the text (Newmark, 1991, p. 83).

2.5.1.2 Grammatical Errors (Syntactic)

Syntactic errors refer to errors that are related to the use of tenses, prepositions, articles as well as the incorrect use of plurality and other parts of speech as well. It is usually related to the structure of the sentence where it can affect the meaning of the output as well. However, it can also not affect the meaning at all. For instance, *"Luka atau bisul yang tidak sembuh, bertukar warna, berdarah atau ulser"* is translated as *"Sores or blisters that do not heal,* ***changes colour,*** *bleeds or ulcers"* (Example from Text 1).

2.5.1.3 Semantic Errors

The errors are related to the meaning of a text. This is when the target text produced incorrect or wrong communicative effect of the source text which directly provides a wrong definition or message. The incorrect meaning from the translated output also results from the incorrect use of lexical items. To illustrate, *"Maklumat peribadi yang berkaitan dengan S Card* ***seperti mata dan token*** *yang dikumpul atau ditebus oleh anda"* is translated as *"Personal information related the S Card as the* ***eyes and tokens*** *collected or redeemed by you"* (Example from Text 4).

2.6 Textual Analysis

According to Halliday, (as cited in Tengku Sepora and Mohamed Abdou, 2006), "a text is a semantic unit" that provides meaning which is produced either in written or oral form. A text may be short or long and it depends on the function of the texts, regardless on how short or long a text is, it still has meaning. Since a text is a semantic unit which provides meaning, therefore, it works at two levels. The two levels are macrostructure and microstructure. In general, at the macrostructure level, it comprises the ideational organisation and the format of a text. Conversely, the microstructure level of a text consists of details such as the sentential and grammatical structure (syntax) and words (lexis).

The textual level (Newmark, 1988) is the base level in translating from the source text into the target text. The grammar of the source text s (clauses and groups) will transpose into the readily equivalents of the target text as well as the lexical units that appear to make sense to the context of sentences. Referential level, on the other hand, is the clarification of all linguistic problems such as words and terms in both

general and specific subjects. Kulwindr Kaur (2006) believes that the substance conveyed from the source text to the target text must be clear and accurate.

Another level, as mentioned by Newmark (1988), is the cohesive level, where he proposes that both structure and moods of the text (article, punctuation marks, linking words, conjunctions – to name a few) are interrelated. This level helps to maintain the coherence and the lengths of paragraphs and sentences of the text. It means, the target text has similar features as in the source text in terms of its style or lengths. In addition, naturalness is the final level of translation. Newmark (1988) states the output of the target text must make sense to the source text and it needs to be read and perceived naturally. Besides that, the grammar for idioms (in the case of literary text) and words are written in ordinary language in order to meet different types of audience or receiver of the target text or language.

It is essential to analyse the four different texts at the microstructure analysis or level. The microstructure analysis involves the lexis, syntax and the meaning of the text (semantic). The rationale of analysing the four texts on the three aspects because all of the elements are interrelated and sometimes it may affect to one another. To illustrate, an excerpt taken from Text 3, *"Dengan berkuatkuasanya Akta Perlindungan Data Peribadi 2010 (**"Akta"**), **ASNB** sebagai pengguna data dikehendaki untuk memastikan **privasi data peribadi pelanggan** yang telah diberikan atau sedang digunapakai atau diselenggara oleh ASNB dipelihara dan dilindungi"* is translated by Google Translate as *"With the Personal Data Protection Act 2010 (**"the act"**), as the **funds** required to **ensure the user data privacy of personal data of customers** who have been granted or are being adopted or are being adopted or maintained by the funds are preserved and protected"*. From the example, the wrong use of lexical word affects the meaning of the sentence as well as due to wrong word order (syntactic) also affects the meaning of the sentence. Newmark (1988, 1991) states that sentence is the basic unit of thought,

presenting an idea or object and affecting the text. Thus, the sentence is the first unit of translation.

In summary, this study attempts at identifying errors on Google Translate output when translating vocative and informative texts (pamphlets and notices). In identifying the errors, error analysis of error categorisation is used as well as textual analysis which analysed the pamphlets and notices at microstructure levels that are related to syntactic, lexical and semantic. Hence, it can be assumed that both of them are interrelated and tied together.

2.7 Translation Quality

A good translation is rendering the original meaning of the source text. The target text output is usually compared to the source text in order to see if the output is according to the norm or violates the meaning and form of the source text. Sager (1983, p. 121) claims there is no standard of translation quality to be exact, as long as it is appropriate and deemed to the purpose of the intended receiver of the target text as he said,

> "There are no absolute standards of translation quality but only more or less appropriate translations for the purpose for which they are intended"
>
> (Sager, 1983, p. 121)

Williams (2004) explains the quality of translation is when the output of the target text can embrace the audience. In regards to the original text, his approach is to analyse if the translated output from the original text is maintained or modified. This enables the output of the target text to be compared and to determine the relevancy of the texts.

A practical overview on translation quality is provided by House (2001, p. 58) who defines that a translation is based on the equivalence of both texts (source text and target text). According to her, there are three types of meaning, namely, semantic, pragmatic and textual. The function of the text depends on the context of a situation where the linguistic notions are classified into syntactic, lexical and textual. She further explicates (2015, p. 38) translation as overt and covert translation where the former does not address the addressees directly while the latter focuses more onto the cultural aspects.

The question arises here is to what extent does machine-aided translation, namely, Google Translate is able to produce a good translation?. The concerns lay on the evaluation of raw output of the target text where Hutchins and Somers (1992) assert that the translation qualities are based on three aspects. There are in terms of accuracy, clarity and style. Accuracy is when the translated output has similar information as in the source text. Clarity concerns with the relevancy of the text among the target text readers and style is related to the appropriateness of the language used. Since this study looks at the errors made by Google Translate in translating vocative and informative texts, there is also a need to look at the comparison between human translations and Google Translate output in order to ensure the quality of translation is achieved.

Human assessment or human post-editing is meaningful and is always needed, but humans are known to be inconsistent and have their own style of interpreting and translating (Graham, 2015, p. 1804). Catford (1965, p. 73) states that when translating a source text into a target text, there are some changes that must be made by human translators due to several factors such as cultural phenomena, linguistic elements or style which are known as translation shift. The end products of the target text are compared in relation to the source text and other human translators to achieve the quality of translation needed.

The use of machine translations pertaining to the source text and target text output is often desired since machine translations is much quicker and is less expensive. However, human judgments or editing is still preferred. Bojar (2011) conducted a study and used four different machine translations and one of them was Google Translate. He first identified the errors found in the machine translation systems and disseminated the output to two human translators. The two human translators needed to rectify the errors found according to the categories of errors listed by Vilar et al. (2006) and judge if the translation was still acceptable.

Popovic and Burchardt (2011) assert that errors produced by machine translation or human can belong to one or more than one categories of error. In fact, there are variations between different human translators when editing the machine translation outputs. Despite the difficulties in assessing the translation output, the error categories identified from the machine translation output helps the human translators to make the required amendments. They believe that the revised version of human translators is still acceptable if it holds the original meaning of the source text although the revised output is based on the machine translation products.

To conclude, to determine whether the translation output which has been rectified by the human translators (for this study) is acceptable or not, it depends on the end product of the target text output. The end product must render the original meaning of the source text.

2.8 Past Studies on Translation and Machine Translation

There are several studies on the use of machine translation in translation, in particular to Google Translate, although most of the studies focused on computational analysis which uses certain software.

One of the studies carried out by Mohamad Nor Amin and Naimah Abdullah (2011) looks at the Arabic and Malay languages and discovers that Google Translate fails to translate some of the nouns, verb phrases and particles from the source text (Arabic) into the target text (BM) correctly. They find out that some of the semantic aspects of the source text cannot be achieved by Google Translate when translating the source text to the target text. Besides that, there are certain words that cannot be translated and needed to be replaced in English. They suggested that experts in Arabic language should provide more Arabic corpus in the MTs in different genres.

Another study conducted by Rensburg et al. (2012) investigates the use of Google Translate in translating the Afrikaans language into English. There are six different texts used in the study (i.e. official letters, newspaper articles, minutes of meetings, class notes, examination notes and PowerPoint slides). The six different texts are first translated by Google Translate from Afrikaans to English and vice versa. Next, the outputs of the source texts are disseminated to two human translators for rectification or editing. One of the translators is a professional translator and the other translator is an English student. There are five raters who assessed the outputs from both of the human translators. The result shows that PowerPoint slides yielded the best result as PowerPoint slides use the simpler sentence structures. Hence, simpler sentences produce better Google Translate output because there is not much rectification to be made by human translators.

Furthermore, a study conducted by Jia et al. (2012) investigate the use of machine translation (MT) to translate the selected titles of Chinese papers from a Chinese journal into English. They construct a website that enabled participants to perform the translation online by using one of the machine translations that is Google Translate. They conclude that at the sentence level of the target text outputs produce by the machine translations, the participants need to spend substantial time for editing. Meanwhile, at the error type level, they identify that machine translations system is unable to provide variation of vocabulary and, thus, produce inaccurate choices of translation.

Condon et al. (2010) study machine translation errors to and from English and Iraqi Arabic and analyse the output at two levels. The first level is done to identify the errors while the second level is to identify how human translators refine the errors. They identify the errors based on Vilar et al.'s (2006) and Keshavarz's (1999) taxonomy of errors and the output then is given to human translators. The outputs later are compared to provide error frequencies. The result shows structural difference between English and Iraqi Arabic in terms of modifier, subject inflection on verbs for Iraqi Arabic, unexpressed pronouns as well as gender categorisation. All of these are due to the linguistic differences which are difficult to translate by machine translation. Hence, human assessment is needed. Some of the errors which have been found do not appear in the machine translation system source or corpora which are found to be difficult for machine translation and other language processing to operate. They provide an example of pronoun error in the context of Iraqi Arabic. The use of 'I' or 'you' does not distinguish between first and second person subjects. They conclude that machine translation fails to achieve one-to-one correspondence between the output and the source text due to linguistic differences which cannot be tackled by the machine translation.

Koponen (2010) conducted a study on error analysis of machine translation in which three English texts were translated into Finnish by two different machine translations and one of them was Google Translate. The selection of texts varies as they contain long and complex sentences as well as short and imperative sentences. The outputs from the machine translations then are compared to human translations in order to show differences between human translations as well as machine translations. In order to analyse the errors made by the machine translations as well as human translation outputs, he applies the unit of translation (UT) or unit of analysis. The analysis of unit of translation is based on the sentence level since it is the largest unit processed by machine translation and it is expected that to have a one-to-one correspondence when translating. The errors identified are classified into five categories. These include omission (where the source text is not conveyed by the target text), addition (the target text word is not present in the source text) and untranslated concept (this is when the source text word cannot be translated and it appears in the target text output). The other two classes are mistranslated concept (where the target text has mistranslated the source text in which the meaning has been distorted) as well as the substituted concept (where there is equivalence of lexical choices from the source text). The results of the error analysis shows that between the two machine translations and one is the statistical system (Google Translate), most errors are found to be in the category of omitting the relation between two concepts as well as adding a new concept to the source text. On the other hand, the outputs produced by human translators reveal that the cases which are identified as omissions and additions differ from outputs produced by machine translations. This is because, human translations have unrelated concepts to the source text compared to machine outputs. This study reveals that the use of error analysis is successful in identifying the differences between machine translations outputs and human translations. However, semantically, the machine

translations outputs still need some corrections. This is especially for long and complex sentences as well as name-like terms where the errors produced by machine more or less affected the meaning of the source text as compared to short sentences.

Koponen and Salmi (2015) study on two short newspaper articles which are translated from English into Finnish by a machine translation, namely Google Translate. Both texts have 673 words with 32 sentences for a long text and 28 sentences for a short text. The machine translation outputs are given to translation students who are majoring in different languages for editing. They classify the errors in terms of its morphological, lexical and syntactic levels. The result shows that the correctness of raw outputs produced by machine translation are not rated highly as 30 sentences out of 120 sentences are translated correctly, others are regard to errors in meaning with language errors. Their study also indicate that long sentences and sentences with a high number of errors are difficult to edit compared to short sentences. They assert that the easy errors include in the word form changes because it can be deduced based on the context in the source text sentence. Meanwhile, the difficult errors are related to missing words or incorrectly translated the words form the source text into the target text, where in some cases it distorts the meaning of the sentence.

Matsuzaki et al. (2016) study the use of machine translation that is involved in a second language proficiency test. 795 high school students are required to answer a few sets of questions that consist of a short English conversation between two people with some of the utterance hidden or left blank forcing the students to choose an appropriate answer. The conversations are translated into Japanese using Google Translate. Based on Vilar et al.'s (2006) framework on error taxonomy, they classify the errors into missing words, word order, incorrect words, incorrect dependency, between two clauses, grammatical property and semantic as well as adding another category which is incorrect dependency. The outputs of Google Translate are then given to the

annotators and they are asked to compare their outputs with machine translation. The comparison of both outputs is made in order to classify the differences according to the errors taxonomy. The study indicates that the output produced by Google Translate is less accurate where some of the clauses translated are irrelevant to the conversation. It also implies here that translations that involve sentences which have ellipsis are difficult to be translated by machine translation as well as human translators or annotators. This is because the context of the conversation is not provided and difficult for human translators to find a correct translation which is relevant prior to the situation in which machine translation, does not have the specialty to do that. Thus, based on their study, the semantic or pragmatic meaning of the text (conversation) could not be rendered as in the original conversation due to some of the semantic elements are missing (i.e ellipsis).

Temnikova (2010) in her study on three different languages such as Bulgarian, Russian and Spanish classifies the errors produced by machine translations based on Vilar et al.'s (2006) classification of error. She identifies the errors and ranks them according to the presumed cognitive effort which is required by human translators to correct them from the easiest to the most difficult error to correct. She states that the easiest errors are related to correct words with incorrect form and followed by the lexical level. For the lexical errors, it involves incorrect synonyms, incorrect words, extra words or missing words. In contrast, the hardest errors is related to the syntactic level which includes wrong or missing punctuation, word order at word level as well as word order at phrase level. She discovers too that long and short sentences are also gives an impact when translated the text especially for sentences that are found to have many errors.

2.9 Conclusion

This chapter looks at the relevant reviews or literature prior to this study. It explains the general terms that are usually used while writing this dissertation especially in analysing the data. The framework that will be used in analysing the data as well as the error identification is presented consistently. It shows that there are several studies have been carried out using both Google Translate and human translators which are relevant to this study. Some of the studies conducted used translation software and certain metrics for analyses.

CHAPTER 3

RESEARCH METHODOLOGY

3.0 Introduction

This chapter will elaborate on the method that is used in examining this study. Therefore, the research design, research framework, research instrument, research subjects of the study, sampling technique, research procedures for data collection and data analysis of this study as well as pilot study will be discussed in detail in this chapter.

3.1 Research Design

The research design employed for this study is a descriptive analysis in nature, investigates the types of errors made by Google Translate and how did the three human translators correct the errors. The qualitative data were gathered in order to look for a phenomenon or pattern (i.e the output of Google Translate). This design is useful for this study as it emphasises on the qualitative data in which the researcher is to identify the patterns obtained from the subjects or samples (Creswell, 2012, p. 245). It also includes human translators' views and perceptions on the use of Google Translate as a channel of translation for expanding the analysis. To obtain their views, a simple set of questionnaire was given to them. The use of questionnaire is to extract human translators' feedback and it should not overlap with the main focus of the study which is qualitative in nature.

3.2 Theoretical Framework

This study aims to examine to what extent does Google Translate help human translators translate from one language to another language particularly from BM to English in terms of its lexical, syntactic and semantic patterns. Thus, the research framework that is relevant prior to this study is based on Keshavarz's (1999) and Vilar et al.'s (2006) framework of errors analysis in translation. The framework is chosen as it underlies the main aspect of the study which looks at the errors made by Google Translate when translating informative and vocative texts (i.e notices and pamphlets). As discussed in Chapter 2 (section 2.5.1), error analysis (EA) is collecting and translating the samples, then identifying and categorising the errors found as well as evaluating and analysing about the errors. The error identification is categorised based on an adapted version of Keshavarz's (1999) error categorisation in terms of syntactic, semantic as well as lexical type which is similar to the taxonomy used by Vilar et al. (2006). Each of the errors is further sub-categorised as follows:

Table 3.1: A Hybrid Model of Error Categorisation (adapted from Keshavarz's (1999) and Vilar et al.'s (2006) framework of error taxonomy)

Errors Category		**Details**
1.	Lexical Errors	- wrong word choices - different word choices
2.	Grammatical (Syntactic) Errors	- tenses - prepositions - articles - plural/singular morpheme(s) - wrong word order -(gerund, conjunction, word form) -subject-verb-agreement (SVA) - word reference
3.	Semantic/Pragmatic Errors	- errors in meaning - collocations

The analysis also cooperated with Newmark's (1988) semantic and communicative approach in translation. Newmark (1988) asserts that when one is translating a text, he or she translates with four (4) levels or less in mind. The four levels are the source text level, the referential level, the cohesive level and the level of naturalness. Translation also involves of three basic processes of interpretation and analysis of the source text as well as the syntactic structure of the source text that corresponds to the target text (Newmark cited in Norizah Ardi et al., 2009).

In order to work with the research framework proposed by Keshavarz (1999) and Vilar et al. (2006) on analysing the errors, a textual analysis is used for this study. Textual analysis functions at macrostructure and microstructure levels where macrostructure deals with the style and format of the text such as written in point or bullet forms, punctuation, markers or overview of the text. On the other hand, the latter focuses on the specific aspects on analysing the text in terms of its semantic, syntactic or lexis components.

Unit of translating (UT) as proposed by Newmark (1988) and Vinay and Darbelnet (1958) is used to identify a unit of translation as a combination of 'lexicological unit' and 'unit of thought' and both can be applied in analysing a text at microstructure analysis or levels which are lexical, syntactical and message. Clearly, unit of translation is used to translate at sentence level or some phrase form of the source text and it is numbered for easier identification. An example is shown in Table 3.2 below:

Table 3.2: Segmentation of unit of translation to identify errors made by Google Translate

Source Text (BM)	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Perubahan dalam *'corak'* pembuangan air besar atau kecil yang berterusan dalam jangkamasa 2-3 bulan.	2	Changes in the *'twist'* defecation or urination continued over a period of 2-3 months.	**Lexical error** - wrong word choice **Semantic error** - error in meaning
Terdapat ketulan yang tumbuh secara tiba-tiba dan tidak hilang diikuti dengan sakit.	4	There are lumps that grow suddenly disappear and not followed by pain.	**Grammatical error** - wrong word order **Semantic error** - error in meaning

Therefore, the theoretical framework used for this study is an adapted version of Keshavarz's (1999) and Vilar et al.'s (2006) error of analysis along with Vinay and Darbelnet's (1958) unit of translation and textual analysis is used to complement this study.

3.3 Research Instruments

This study used four (4) samples of different texts produced by Google Translate in order to collect the data needed. The outputs of Google Translate from BM to English were given to three (3) human translators for editing. The human translators edited the target text based on the Google Translate output and then produced the final copy of the target text. Additionally, the researcher conducted a phone interview with

the human translators. The interview was conducted in order to obtain additional information on their practice in using Google Translate and whether any additional tool is used in their translation job. Due to constraints of time and different locations between the researcher and participants, a phone interview was conducive to conduct although it was costly (Creswell, 2012). Apparently, the phone interview was conducted to ask similar questions used in the questionnaire (see Appendix E). The researcher also provided a simple set of questionnaire (Appendix E) which consisted of open-ended questions and a few questions that employed the Likert scale to the three human translators. Again, the relevance of giving the questionnaire to them is to determine their perceptions of using Google Translate in translation and their views on the translation quality as well.

3.4 Research Data

The researcher has taken into consideration documents that are usually written in bilingual languages (BM and English) and received by the target texts' readers. Therefore, notices and pamphlets were chosen for this study. The four texts were:

i) Text 1: *10 Tanda-tanda Amaran Kanser*

ii) Text 2: *Persatuan Kebangsaan Kanser Malaysia*

iii) Text 3: *Akta Perlindungan Data Peribadi (2010) - Polisi Privasi (Amanah Saham Nasional Berhad)*

iv) Text 4: *SOGO Kuala Lumpur Notis Privasi*

The notices are taken from *SOGO Kuala Lumpur Deparment Store Sdn. Bhd Privacy Notice* issued in 2014 and *Amanah Saham Nasional Berhad* (ASNB). The former notice has approximately 550 words with 52 units of translations (UTs). The notice consists of some phrases while some of the information is written in point form. The purpose of the notice is to inform the customers' rights and personal information while the latter notice urges the policy holder of ASNB to be always updated and ensure the policy is still in use. On the other hand, the second notice has approximately 450 words and had 18 units of translations. This notice also has information that is written in point form.

Pamphlets produced by *Persatuan Kebangsaan Kanser Malaysia (National Cancer Society Malaysia)* on early signs of cancer as well as holistic treatments for cancer were chosen. Both of the pamphlets have 215 words and 210 words respectively where the first pamphlet on early signs of cancer had 24 units of translations while the second pamphlet on holistic treatments for cancer had 18 units of translations. Similarly to notices, the pamphlets are also written in point form. It can be seen that notices and pamphlets vary based on their lengths where notices are longer than pamphlets but both shared the same features in which some of the sentences are written in point form.

The samples for this study are chosen on the basis of giving and providing information to the readers from different sources that are reachable by them which used BM and English, as bilingual versions are always needed (Norhazlina Husin, 2009). As for the notice policy, it is a customary act for the policy holder to renew his or her policy every year. On the other hand, the pamphlets on cancer were chosen due to health facts to ensure that everyone receives and understands the information conveyed regardless of gender and races. Hidayatul Akmal (2007) reports that approximately 4,000 new cases on cancer are diagnosed in Malaysia and colon cancer is said to increase every year and is the most common cancer that attacks men apart from breast

cancer. Thus, having bilingual languages or versions (BM and English) for this information is really useful. Likewise, Lauer (1995) asserts that medical brochures or pamphlets in English and Spanish are frequently found in health centres for transmitting specialised information to the public.

Besides that, the research tool used for this study is Google Translate. Google Translate is used in the first place to provide a quick translation of the target text (English). Henry (2014) conducted a survey and found that the use of Google Translate for translation has the highest percentage (45%) as compared to other translation tools. According to him, since it can be used on a smart phone which is easy and provides quick translation, thus, making it as one of the popular translation tools among the users, especially the second language learners. Hence, Google Translate is chosen as a subject research tool for this study because it is free and available for everyone at any time via online.

3.5 Participants

In obtaining reliable participants, the researcher chose three female freelance translators who have experience in translating technical and general texts in both languages (BM and English). All the three female freelance translators graduated from Universiti Sains Malaysia (USM). Universiti Sains Malaysia is one of the local universities that offers courses in translation and interpreting in which the participants are familiar with the methods and strategies in translating and interpreting, legal translation, translation evaluation as well as computer-aided translation. The backgrounds of the participants are explained as follows:

Participant 1: Graduated in Bachelor of Arts in Translation and Interpreting (2009) and Masters in Translation (2014). The participant has more than four years of experience in translation. Among her works in translation are *Descriptive Translation Studies: And Beyond* which is a book, translated questionnaires for a medical college, works in various technical texts, and handled terminology for translation and reviewed questionnaires for Microsoft projects.

Participant 2: Graduated in Bachelor of Arts in Translation and Interpreting (2012). The participant has three years of experience in translating and editing. She is the founder of *ProofLaju* and currently works as a content editor for a company named *IHS* in Penang. Her translation works are related to journal articles, thesis for Masters and PhDs as well as technical texts.

Participant 3: Graduated in Bachelor of Arts in Translation and Interpreting (2013) and currently pursuing Master's in Translation in Universiti Sains Malaysia (USM). The participant has approximately three years of experience in translating. Some of her translation works are technical texts, academic texts and websites.

3.6 Research Procedures

The research procedures for this study were conducted accordingly in order to achieve the objectives and purpose of this study. Firstly, the researcher identified different texts that were written in bilingual forms (i.e notices and pamphlets). Next, the researcher inserted the texts on Google Translate in order to obtain a quick translation

output. The texts were inserted according to their formats or preferences as even some of the sentences had subheadings. The outputs produced by Google Translate were then marked sentence by sentence for each of the texts based on Vinay and Darbelnet's (1958) unit of translation in which some of them were in bullet points. The unit of translation was numbered from 1 onwards.

Once the texts were literally translated by Google Translate, the researcher identified the errors produced by Google Translate based on an adapted taxonomy of error categories proposed by Keshavarz's (1999) and Vilar et al.'s (2006) framework. Next, the raw outputs from Google Translate were then disseminated to the subjects (human translators) via e-mail for editing. The three human translators were not provided with the errors as they needed to rectify the Google Translate outputs based on the source texts given. The three human translators were told about the objectives of this study and the researcher's intended outcome instead. Additionally, due to other commitments faced by human translators, the time frame given for them to complete the tasks was within two weeks. Each of the human translators was paid although they were just required to rectify the Google Translate outputs and some of the outputs were not corrected as it was relevant and accepted in the source text. Moreover, the researcher conducted phone interviews with the three human translators to obtain their perceptions and opinions on using Google Translate upon the completion of their tasks.

Therefore, the procedures of data collection for this study were conducted in two phases:

i) Identified bilingual texts and translated them by using Google Translate.

ii) Disseminated the raw outputs of Google Translate to human translators for editing and conducted phone interviews with them to obtain their perceptions.

3.7 Summary of Data Analysis Procedures: Conceptual Framework

Figure 3.1 below shows the data analysis procedures applied in the study.

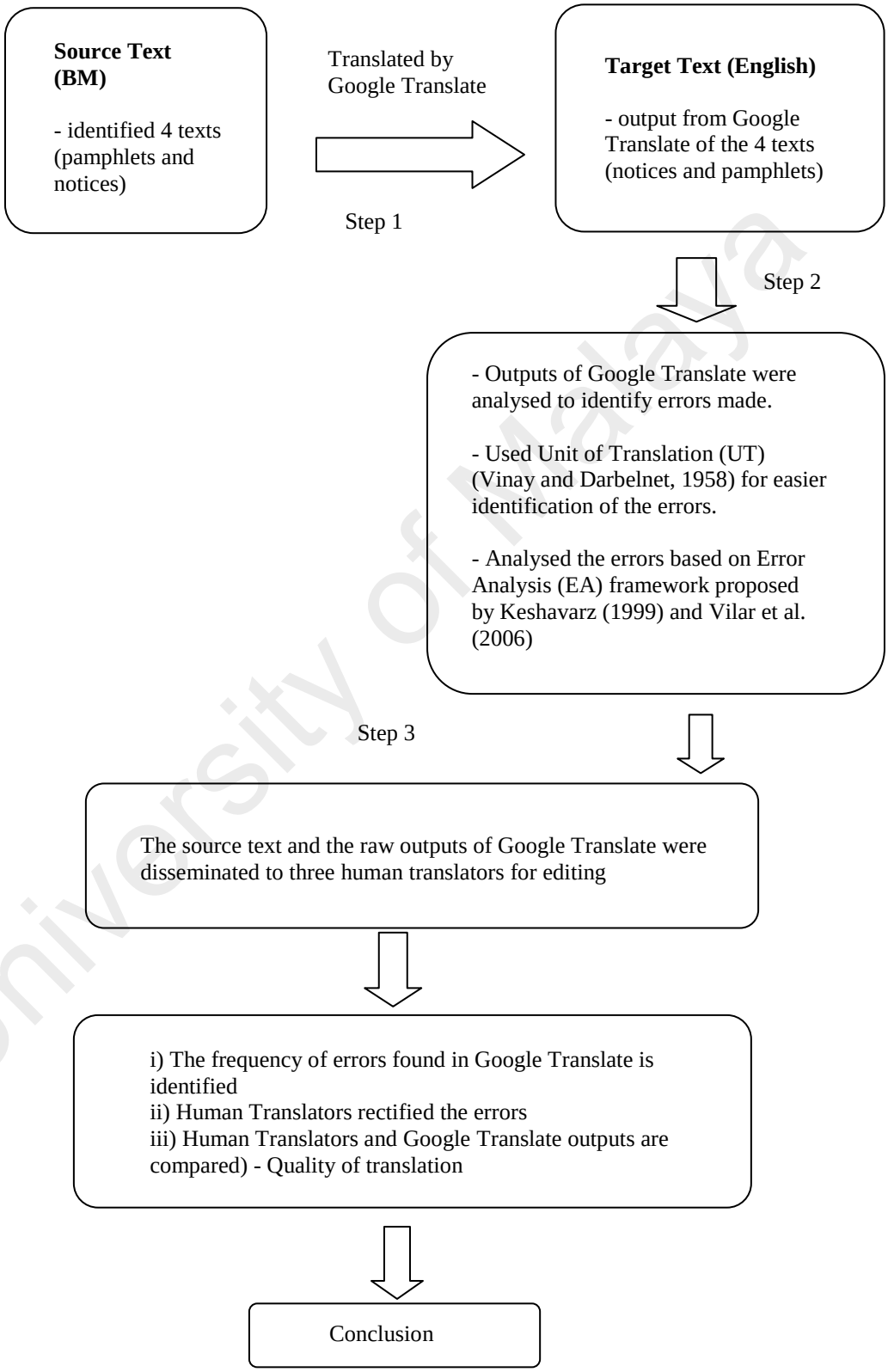

Figure 3.1: Summary of analysis of the study

3.8 Conclusion

This chapter discusses the methodology used in conducting the research in which it involved different phases and procedures to collect the data. Besides that, the conceptual framework is visualised for a better overview of this research as well as the theoretical framework is discussed and explained in an in-depth manner. Due to the fact that this study involved analysing the translation output from Google Translate, the reliability and validity of the data procedures and collection are carefully selected. The researcher has ensured that the participants (human translators) for this study are those who have knowledge in the translation area although their methods or strategies used in translation are different.

CHAPTER 4

FINDINGS AND DISCUSSIONS

4.0 Introduction

This chapter presents and discusses the results obtained from the findings of the four texts (notices and pamphlets). The findings were analysed and presented in tables according to the proposed research questions. The four texts were translated using Google Translate and each of the texts was tabulated according to Vinay and Darbelnet's (1958) unit of translation for easier identification. The errors found were categorised based on Keshavarz's (1999) and Vilar et al.'s (2006) error analysis framework.

4.1 Text Format

The textual analysis of the texts were examined in terms of its macro structure or macro textual elements such as the purpose of the text or the style used in the texts before being examined at the microstructure levels. Koponen (2010) states that macro textual analysis is focused on the purpose, function or effect of a text. Based on the four texts (notices and pamphlets), both types serve similar functions: vocative and informative. Its purpose is to inform the readers or audience of the texts to do or to take certain actions as deemed by the texts. Both notices and pamphlets were written in point form although notices were longer than pamphlets. When they were translated, the overall meaning of the target texts output still render the original meaning of the source texts. This means, the translation produced by Google Translate still retains the meaning of the original text.

The format for both source texts and target texts were preserved where Google Translate maintained the original layout of the source texts when translated online. For example in Text 3: *Akta Perlindungan Data Peribadi (2010)- Polisi Privasi (Amanah Saham Nasional Berhad),* in unit translation 2, "*AKTA PERLINDUNGAN DATA PERIBADI 2010-POLISI PRIVASI"* was translated in capital letters as in the source text *"PERSONAL DATA PROTECTION ACT 2010- PRIVACY POLICY"*. Apart from that, Google Translate retained the style of the source text by keeping the similar visual format of the target text where some of the notices and pamphlets were divided into subheadings and numbered. Nonetheless, there were inconsistencies in spelling as in Text 1: *10 Tanda-tanda Amaran Kanser,* the use of 'centre' or 'center' was identified throughout the target text output, which was translated by Google Translate, although both were acceptable.

Hence, the textual analysis for both source and target texts were alike where Google Translate was able to keep and maintain the original format of the source text. It can be said that Google Translate did not encounter any problem in relation to text format or style.

4.2 Error Categorisation

As discussed in sections 2.5.1 and 3.2, the error categorisation was based on Keshavarz's (1999) and Vilar et al.'s (2006) taxonomy of error analysis. The adapted model proposed by them was useful to answer research question 1 which was to identify the typical errors made by Google Translate when translating vocative and informative texts. The hybrid version of the errors was classified into three errors respectively which were Lexical Errors, Grammatical Errors (syntactic) and Semantic Errors. Each of the types of errors consisted of sub-errors (refer to Table 3.1, p. 36). The four texts (notices

and pamphlets) were tabulated into tables and marked at a sentence or phrase level by using a unit of translation where only errors found in Google Translate outputs were identified.

The categorisation of the errors helps to identify the typical errors made by Google Translate and the frequency of that particular error. This categorisation of errors was helpful in finding which category among the three, namely lexical, grammatical (syntactic) and semantic had the highest frequency of errors made and encountered by Google Translate when translating notices and pamphlets from BM to English. To reiterate, only the errors made by Google Translate were analysed along with their units of translation. The correct translations produced by Google Translate were not shown.

4.2.1 Text 1: *10 Tanda-tanda Amaran Kanser*

Text 1 is a pamphlet from a medical perspective. It urges the readers to be aware of the signs of cancer particularly for women. The text consistsed of approximately 215 words and had 24 units of translation where some were written in point form.

Table 4.1: Errors made by Google Translate in Text 1 (Pamphlet)

Source Text (BM) **(Pamphlet)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Perubahan dalam *'corak'* pembuangan air besar atau kecil yang berterusan dalam jangkamasa 2-3 bulan.	2	Changes in the *'twist'* defecation or urination continued over a period of 2-3 months.	**Lexical error** - wrong word choice **Semantic error** - error in meaning

Terdapat ketulan yang *tumbuh secara tiba-tiba* dan *tidak hilang diikuti dengan sakit.*	4	There are lumps that grow *suddenly disappear* and *not followed by pain*.	**Grammatical error** - wrong word order **Semantic error** - error in meaning
Luka atau bisul yang tidak sembuh, *bertukar warna*, berdarah atau ulser.	7	Sores or blisters that do not heal, *changes color*, bleeds or ulcers.	**Grammatical error** - tenses
Senak atau *kesukaran untuk menelan.*	9	Indigestion or *difficulty swallowing*.	**Grammatical error** - missing preposition
Suara menjadi serak atau *batuk melebihi 2 minggu* terutama pada mereka yang tidak merokok.	10	*Voice become hoarse* or cough for more than 2 weeks, especially in those who do not smoke.	**Grammatical error** - subject-verb-agreement (SVA)
Batuk penghisap rokok yang bertambah teruk.	11	*Smokers cough* that *worsens*.	**Grammatical error** - missing possessive marker -SVA
Untuk jenis kanser *yang boleh* dikesan awal, *terdapat 80%* peluang untuk terus hidup.	15	For *this type of cancer can be detected early*, there is *a 80%* chance of survival.	**Grammatical error** - missing relative pronoun - wrong article
Mengekalkan berat badan unggul	20	Maintaining *ideal body* weight	**Grammatical error** - missing article

Based on the above table (Table 4.1), the errors were categorised into three categories, namely, lexical errors, grammatical errors and semantic errors. From the analysis of error identification, grammatical errors had the highest frequency with nine (9) errors, two (2) for semantic errors and one (1) for lexical errors. It is noted that some of the errors can be categorised into more than one type of error.

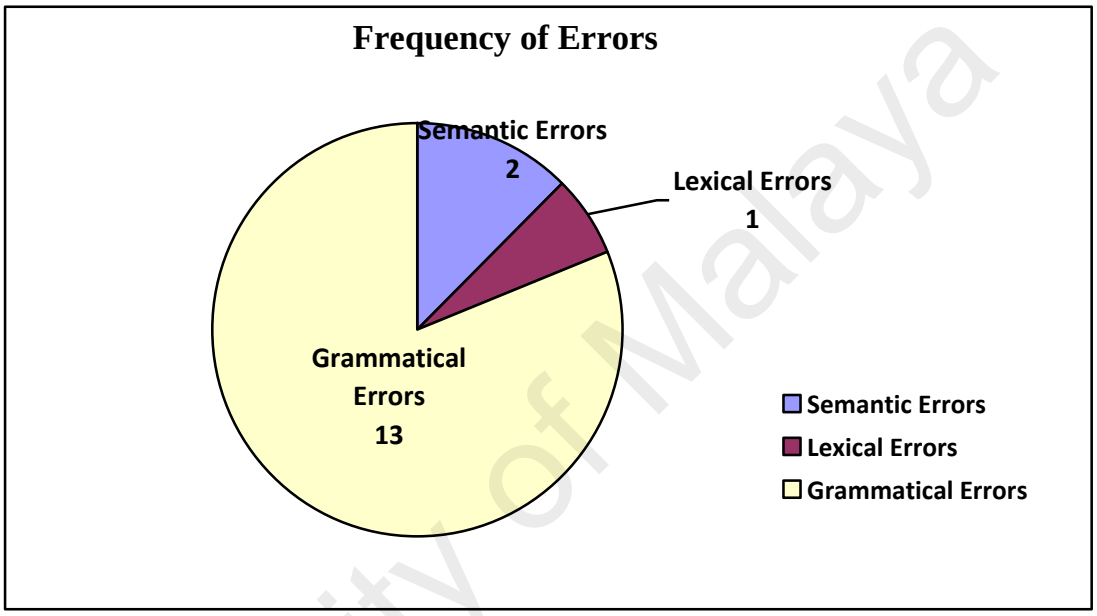

Figure 4.1: Frequency of Errors in Text 1 (Pamphlet)

Each of the categories of error was further divided into its specific error type. Based on Figure 4.1, grammatical errors were found to have the highest number of errors. In this grammatical category, the specific errors found were wrong word order, wrong tense use, missing preposition, missing possessive marker, missing relative pronoun, wrong article use as well as missing article where each of these errors was rated with one error. Since grammatical errors have more specific errors under its category compared to lexical and sematic errors, the tendency of Google Translate to translate wrongly was higher based on the unit of translation applied. Hence, it was found that the error made by Google Translate in most of the unit of translation was in terms of the grammatical category. Meanwhile, incorrect use of subject-verb-agreement

(SVA) rated as two (2) errors. There were two (2) semantic errors where the errors were in meaning due to wrong word choices that affected the meaning of the sentence. For lexical errors, it was identified that there was only one (1) error due to a wrong word choice used which also resulted in errors in meaning.

Table 4.2: Types of Error in Text 1 (Pamphlet)

Errors Category	**Types of errors**	**Frequency**
Lexical Errors	Wrong word choice	1
Grammatical Errors	Wrong tenses	1
	Missing preposition	1
	Missing possessive marker	1
	Missing relative pronoun	1
	Wrong article	1
	Missing article	1
	Subject-verb-agreement (SVA)	2
	Wrong word order	1
Semantic Errors	Errors in meaning	2

Thus, the most common types of errors found in Text 1 were the use of articles and errors in meaning and subject-verb-agreement (SVA). The meaning was affected by the wrong lexical choice as well as wrong word order where it distorted the meaning of the sentences.

4.2.2 Text 2: *Persatuan Kebangsaan Kanser Malaysia*

Text 2 is similar to Text 1 where it is taken from the same resource. It provides information regarding the sources and treatments for cancers. It consistsed of 18 units of

translations with an approximate of 210 words. Similar to Text 1, some of the texts were written in point forms and had a sub-categorisation under each heading.

Table 4.3: Errors made by Google Translate in Text 2 (Pamphlet)

Source Text (BM) **(Pamphlet)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Sila simpan risalah ini atau berikan kepada yang tersayang.	2	Please keep this brochure or *give to* a loved one.	**Grammatical error** - missing object pronoun
Pusat Kebangsaan Kanser Malaysia (NCSM) adalah satu-satunya badan kebajikan yang menawarkan perkhidmatan rawatan kanser dan *sokongan yang menyeluruh kepada para pesakit dan ahli-ahli sokongan mereka* melalui pusat-pusat kanser NCSM.	3	National Cancer Centre Malaysia (NCSM) is the only national charity that offers cancer treatment services and comprehensive support to patients and *members of their support* through NCSM cancer centers.	**Lexical error** - wrong word choice **Semantic error** - error in meaning
Pusat Sumber dan *Kesejahteraan*	4	Resource Centre and *Welfare*	**Lexical error** - wrong word choice **Grammatical error** - wrong word order
Diurus oleh seorang kaunselor berpengalaman dan para sukarelawan untuk menjawab pertanyaan awam.	6	*Managed by an experienced counselors* and volunteers to answer public inquiries.	**Grammatical error** - SVA

Table 4.3: Errors made by Google Translate in Text 2 (Pamphlet) (continued)

Source Text (BM) **(Pamphlet)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Pusat ini juga mempunyai sebuah perpustakaan rujukan kanser dan *sering menganjurkan mesyuarat kumpulan-kumpulan sokongan*, bengkel dan kelas *kesejahteraan* seperti yoga dan qigong.	7	The center also has a reference library and often *organizes meetings cancer support groups*, workshops and classes such as yoga and qigong *welfare.*	**Grammatical error** - missing preposition - wrong word order
Pakar perubatan *di pusat ini* memberi perkhidmatan pengesanan kanser untuk wanita daripada setiap peringkat umur.	9	Physicians at *these centers* provide cancer detection services for women of all ages.	**Semantic error** - error in meaning (reference is made to a specific center)
Perkhidmatan termasuk mamografi digital, ultrasound, *ujian pangkal rahim*, pemeriksaan payudara sendiri dan cara pemakaian payudara palsu bagi pesakit yang baru menjalani pembedahan payudara.	10	Services include digital mammography, ultrasound, cervical exams, breast self-examination and how to use *fake breasts* for patients who have recently undergone breast surgery.	**Semantic error** - collocation
Pasukan pakar teknologi perubatan kami akan mengimbas dan memberi pengesanan awal tentang penularan (metastasis) kanser *di organ badan dan tulang*.	13	Our team of experts in medical technology will scan and provide early detection of infection (metastasis) of cancer *in the body organs and bones.*	**Grammatical error** - wrong word order

Table 4.3: Errors made by Google Translate in Text 2 (Pamphlet) (continued)

Source Text (BM) **(Pamphlet)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Rawatan untuk pesakit-pesakit kanser termasuk radioterapi, kemoterapi, *terapi hormon* dan brakiterapi kadar dos tinggi.	15	Treatment for cancer patients *including* radiotherapy, chemotherapy, hormonal therapy and high dose rate brachytherapy.	**Grammatical error** - wrong word form
Pusat ini dilengkapi dengan peralatan berteknologi tinggi *setaraf dengan hospital swasta* tetapi *dengan caj yang lebih murah* dan sedikit atau *tiada masa menunggu.*	16	The center is equipped with high-tech equipment *equivalent to the private hospital* but with *charges cheaper* and little or *no wait time.*	**Grammatical error** - wrong word order - gerund
Rumah Harapan Kanak-kanak	17	*Hope childhood home*	**Lexical error** -wrong word choice **Grammatical error** - wrong word order
Berkonsepkan ‘rumah kedua’, NCSM menyediakan tempat penginapan bagi pesakit kanak-kanak luar bandar dan penjaga mereka yang menjalani rawatan di Kuala Lumpur.	18	*Just as ‘second home’ NCSM provide* accommodation for *patients of rural children* and their caregivers who are undergoing treatment in Kuala Lumpur.	**Grammatical error** - missing article - SVA -wrong word order

Based on Table 4.3 above, it was found that grammatical errors had the highest frequency of errors that was rated with thirteen (13) errors. Lexical errors occured three times similar to semantic errors.

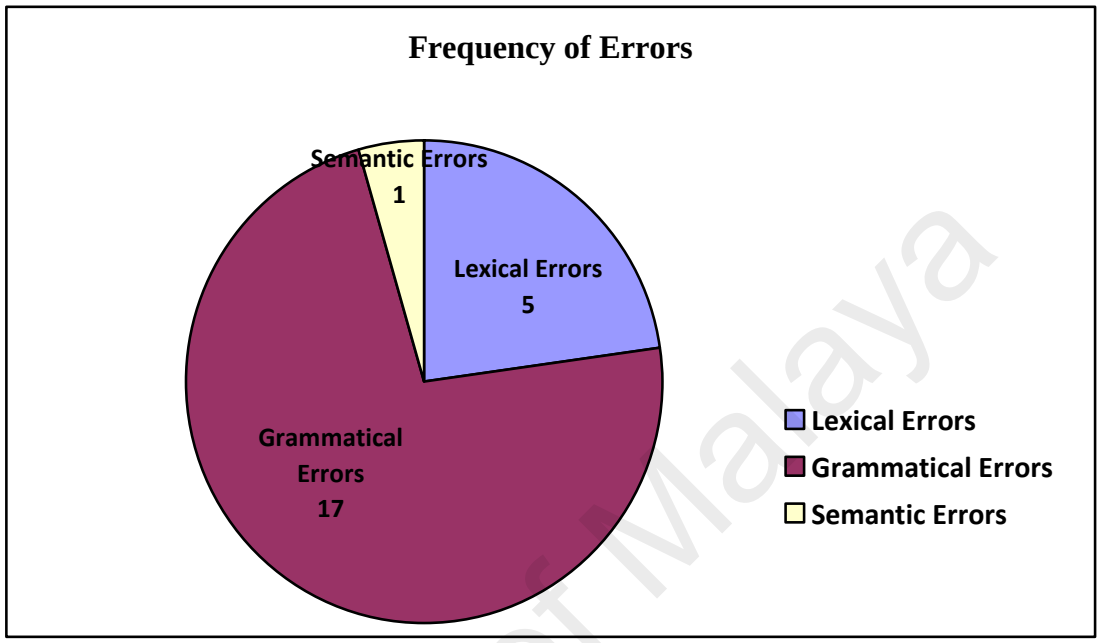

Figure 4.2: Frequency of Errors in Text 2 (Pamphlet)

The analysis of errors was further broken down into error types for each category where wrong word order showed a higher number of errors: six (6), which falls under grammatical errors. The second number of highest errors made by Google Translate in Text 2 was wrong word choices which occurred three (3) times. The incorrect use of subject-verb-agreement (SVA) and errors in meaning had the same frequency that was two (2). Other errors such as missing object pronoun, missing preposition, missing article, gerund as well as collocation from both the grammatical and semantic error categories had an equal number of frequency that was 1 (one) for each type based on the unit of translation. Similar to Text 1, in Text 2, Google Translate encountered a problem when translating a simple and small unit of word such as preposition or article, as it was either translated wrongly or not translated at all.

Table 4.4: Types of Error in Text 2 (Pamphlet)

Errors Category	**Types of Errors**	**Frequency**
Lexical Errors	Wrong word choice	3
Grammatical Errors	Missing object pronoun	1
	Missing preposition	1
	Missing article	1
	Wrong word order	6
	Subject-verb-agreement (SVA)	2
	Gerund	1
	Wrong word form	1
Semantic Errors	Error in meaning	2
	Collocation	1

As shown in Table 4.4 above, wrong word order was highly recorded for Text 2 in which Google Translate had a problem when translating some of the sentences. It was noticeable that Google Translate encountered the problem in regards to the use of article, subject-verb-agreement (SVA), prepositions which similarly were found in Text 1. Besides that, wrong word choices were also one of the most common types of error made by Google Translate.

4.2.3 Text 3: *Akta Perlindungan Data Peribadi 2010 - Polisi Privasi (Amanah Saham Nasional Berhad)*

Text 3 is a notice from *Amanah Saham Nasional Berhad (ASNB).* As a vocative and informative text, it informs the users or holders of the policy to renew their ASNB policies. The text had 450 words with 18 units of translations. The words were longer than the two previous pamphlets and some of them were written in bullet points and had sub-division under each heading.

Table 4.5: Errors made by Google Translate in Text 3 (Notice)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
1. *Pihak kami mengucapkan terima kasih* kerana menjadi pelanggan setia Amanah Saham Nasional Berhad (*"ASNB")*	3	1. *The wishes we thank you* for being a loyal customer Amanah Saham Nasional Berhad (*"funds"*).	**Lexical error** - wrong word choice **Grammatical error** - missing preposition
2 .Dengan berkuatkuasanya Akta Perlindungan Data Peribadi 2010 ("Akta"), *ASNB* sebagai pengguna data dikehendaki untuk memastikan privasi data peribadi pelanggan yang telah diberikan atau sedang digunapakai atau diselenggara oleh ASNB dipelihara dan dilindungi.	4	2. With the Personal Data Protection Act 2010 *("the act"), as the funds required to ensure the user data privacy of personal data of customers who have been granted or are being adopted or maintained by the funds are preserved and protected.*	**Lexical error** - wrong word choice **Grammatical error** - wrong word order - SVA **Semantic error** - error in meaning
3. Justeru itu, kami ingin memaklumkan bahawa data peribadi dan segala maklumat anda yang telah diberikan kepada *ASNB* semasa memohon untuk menjadi *pemegang unit* bagi *tabung unit amanah* yang diuruskan oleh *ASNB* ("Data Peribadi"), akan digunakan dan diproses oleh ASNB bagi tujuan menyediakan produk dan perkhidmatan kami serta perkhidmatan berkaitan yang lain.	5	3. Therefore, we would like to *inform you the personal data* and any information which you have given to the *funds* when applying to become *a holder of units* in a *unit trust fund* managed by the *funds* ("Personal Data"), will be used and processed by the funds for the purpose of providing products and services and other related services.	**Lexical error** - wrong word choice **Grammatical error** - SVA - missing relative pronoun

Table 4.5: Errors made by Google Translate in Text 3 (Notice) (continued)

Source Text (BM) **(Notice 3)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
4. Sebagai pemegang unit *ASNB*, anda berakujanji, mengakui dan bersetuju bahawa ASNB boleh:	6	As the Holders of the *funds*, *you undertakes, acknowledge and agrees* that the funds can be:	**Lexical error** - wrong word choice **Grammatical error** - SVA
a) Mengumpul, menyimpan, memproses dan menggunakan Data Peribadi anda selaras dengan Polisi Privasi *ASNB* seperti yang dinyatakan dalam laman web kami di www.asnb.com.my; dan	7	a) Collect, store, process and use your Personal Data in accordance with the Privacy Policy *funds* as set out in our website www.asnb.com.my; and	**Lexical error** - wrong word choice **Grammatical error** - missing preposition
b) *Menzahirkan* dan berkongsi Data Peribadi anda *di kalangan kumpulan syarikat*, ejen yang diberi kuasa dan pembekal perkhidmatan (yang telah bersetuju untuk menyimpan Data Peribadi tersebut secara sulit) bagi tujuan menyediakan produk dan perkhidmatan berkaitan yang lain dan bagi tujuan lain.	8	b) *Demonstrate* and share your Personal Data *among group companies*, authorized agents and service providers (who agreed to keep the Personal Data confidential) for the purpose of providing products and other related services, and for other purposes.	**Lexical error** - different word choice **Grammatical error** - wrong word order
5. Adalah menjadi satu kewajiban mandatori untuk anda memberikan Data Peribadi kepada kami.	9	5. It shall be a mandatory obligation for you to provide Personal Data to us.	**Grammatical error** - missing possessive pronoun

Table 4.5: Errors made by Google Translate in Text 3 (Notice) (continued)

Source Text (BM) **(Notice 3)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Jika anda *enggan memberikan* dan tidak membenarkan Data Peribadi anda digunakan bagi tujuan-tujuan yang telah dinyatakan, ia boleh menghalang anda untuk melanggan atau menggunakan sepenuhnya produk dan perkhidmatan kami yang mana pihak kami tidak bertanggungjawab dan tidak boleh dipertanggungjawabkan sekiranya perkara tersebut berlaku.	10	If you *refuse to leave* and *do not allow use* of your Personal Data for the purposes that have been disclosed, it may prevent you *to subscribe or use* of our products and services which we are not *responsible and* cannot be held responsible *if the event occurred*.	**Lexical error** - wrong word choice **Grammatical error** - missing preposition - tenses - wrong word order
6. Anda berhak untuk mengakses Data Peribadi anda *yang disimpan* oleh kami dan membuat permintaan untuk mengemaskini *atau membuat pembetulan Data Peribadi yang tidak tepat* atau mengelirukan atau menghadkan pemprosesan Data Peribadi dengan menghantar permohonan secara bertulis kepada alamat dinyatakan di bawah:	11	6. You have the right to access your Personal Data *held* by us and make a request to update or otherwise *correct the Personal Data is inaccurate* or misleading or limiting the processing of personal data by sending *the* written application to the address indicated below:	**Lexical error** - wrong word choice **Grammatical error** - missing relative pronoun - wrong article

Table 4.5: Errors made by Google Translate in Text 3 (Notice) (continued)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
7. Jika anda tidak bersetuju untuk nama anda dan *butir-butir yang berkaitan* digunakan oleh kami bagi tujuan penghantaraan apa-apa bahan pemasaran atau anda enggan dihubungi oleh kami bagi apa-apa penjualan dan promosi sebarang produk atau perkhidmatan, sila maklumkan secara bertulis kepada alamat berikut:	12	7. If you do not agree to your name and *particulars pertaining* used by us for the purpose of sending any marketing material or you refuse to be contacted by us for any sales and promotions of any products or services, please notify us in writing to the following address:	**Grammatical error** - wrong word order
Amanah Saham Nasional Berhad Jabatan Perhubungan Pelanggan & Ejen UG, *Balai PNB*, 201-A Jalan Tun Razak 50400 Kuala Lumpur	13	Amanah Saham Nasional Berhad Department of Customer Relations & Agents UG, *the PNB*, 201-A Jalan Tun Razak 50400 Kuala Lumpur	**Lexical error** - different word choice
Atau, anda boleh menghantar e-mel ke asnbcare@pnb.com.my	14	Or, you can *send e-mail* asnbcare@pnb.com.my	**Grammatical error** - missing article -missing preposition
Sekiranya kami tidak menerima sebarang maklum balas daripada anda berhubung dengan penggunaan Data Peribadi anda dalam tempoh 14 hari dari tarikh notis ini, anda dianggap bersetuju dengan terma dan syarat yang dinyatakan di sini.	16	If we do not receive any feedback from you regarding the use of your Personal Data within 14 days of the date of this notice, you are deemed to agree to terms and conditions set forth herein.	**Grammatical error** -missing article

As shown in Table 4.5, grammatical errors were found to be the most frequently made errors by Google Translate when translating the text. In total there were twenty-eight (28) errors, consisted of eighteen (18) errors for grammatical category, lexical errors were nine (9) in total followed by only one (1) error for semantic errors.

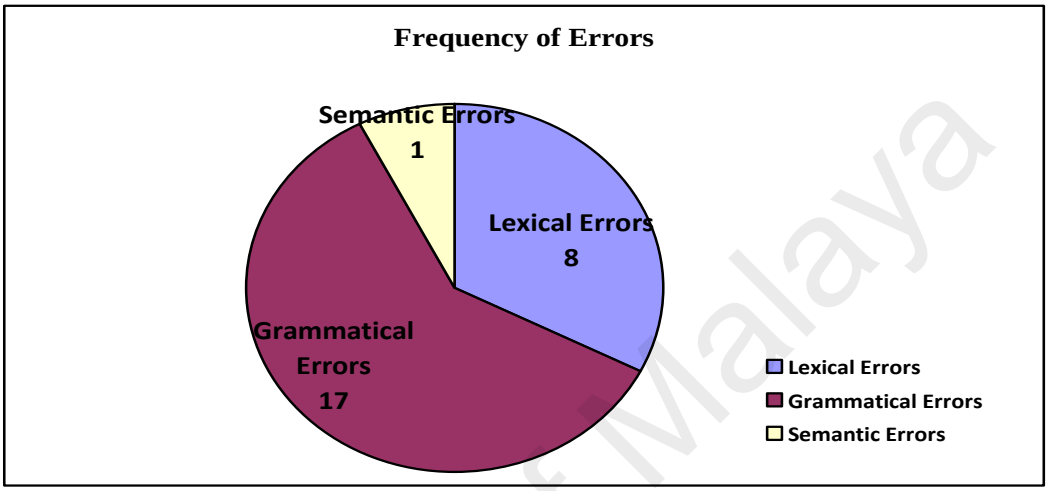

Figure 4.3: Frequency of Errors in Text 3 (Notice)

Further analysis on each category of errors resulted in wrong word choices, under lexical errors category were identified as the most frequent error made by Google Translate. Besides that, the analysis also showed that missing prepositions and wrong word order shared a similar frequency that was four (4). It was followed by missing pronouns (relative and possessive) and subject-verb-agreement (SVA) where both occured three (3) times. The analysis also revealed that different word choices had two (2) frequencies of errors which were similar to missing articles for grammatical category. The use of wrong tenses, wrong article as well as error in meaning (semantic errors) had one (1) error for each.

Table 4.6: Types of Error in Text 3 (Notice)

Errors Category	Types of Errors	Frequency
Lexical Errors	Wrong word choice	7
	Different word choice	2
Grammatical Errors	Missing preposition	4
	Wrong word order	4
	Subject-verb-agreement (SVA)	3
	Missing relative pronoun	2
	Missing possessive pronoun	1
	Missing article	2
	Wrong article	1
	Tenses	1
Semantic Errors	Errors in meaning	1

Therefore, the typical types of errors which were identified in Text 3 (Notice) were wrong word choices followed by missing preposition and wrong word order as both were noted to have the highest frequency after wrong word choices.

4.2.4 Text 4: *SOGO Kuala Lumpus Notis Privasi*

Text 4 is a notice taken from *SOGO Kuala Lumpur* of Privacy Notice. Similar to Text 3, it also informs the holders or customers regarding the policy of the company on the use of their personal information for the company or department reference. The text had 550 words and 52 units of translations. Similar to Text 3, Text 4 too was compiled in point forms where it had its own heading for each.

Table 4.7: Errors made by Google Translate in Text 4 (Notice)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Notis Privasi ini adalah untuk memaklumkan anda mengenai polisi SOGO (K.L) Department Store Sdn.Bhd *(syarikat/kami)* mengenai polisi untuk mengumpul, menggunakan, memproses dan mendedahkan Maklumat Peribadi anda.	1	This privacy notice is to inform you about the *policies SOGO* (K.L) Department Store Sdn.Bhd *(a/us*) *on policies* for collecting, using, processing and *unwind disclose your* Personal information.	**Lexical error** - wrong word choice **Grammatical error** - gerund - word order
Sila ambil perhatian bahawa Notis Privasi ini adalah berkaitan dengan pelbagai sumber yang mengumpul Maklumat Peribadi anda termasuk tetapi tidak terhad kepada *pelanggan-pelanggan, pelawat-pelawat, penyewa-penyewa dan pembekal-pembekal kami*.	3	Please note that this Privacy Notice is associated with *a variety of resources that collect* your Personal Information including but not limited to, *customers, visitors, tenants and our suppliers.*	**Grammatical error** - wrong word order **Semantic error** - collocation
1.1 Maklumat Peribadi yang kami kumpul mengenai anda termasuk tetapi tidak terhad kepada maklumat berikut:-	5	1.1 Personal information we collect about you including but not limited to the following information:-	**Grammatical error** - missing relative pronoun
Maklumat peribadi yang berkaitan dengan S Card *seperti mata dan token* yang dikumpul atau ditebus oleh anda;	7	Personal information related to the S Card *as the eyes and tokens* collected or redeemed by you;	**Lexical error** - wrong word choice **Semantic error** -collocation

Table 4.7: Errors made by Google Translate in Text 4 (Notice) (continued)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Maklumat lain yang telah atau mungkin dikumpulkan oleh kami atau yang anda berikan kepada kami dari masa ke semasa berkaitan dengan apa-apa perkhidmatan, transaksi, pertandingan, kajian, promosi, *soal selidik* atau *komunikasi dengan kami* yang berkaitan dengan produk kami atau selainnya;	10	Other information that has been or may be collected by us that you provide to us from time to time *in connection with any service, transaction, contests, surveys*, promotions, *surveys or communicate with us* in connection with our products or otherwise;	**Lexical error** - wrong word choice **Grammatical error** - wrong word part
Imej anda *yang* telah dirakam oleh *video pengawasan* yang dipasang di beberapa bahagian premis kami yang merupakan sebahagian daripada infrastruktur keselamatan kami; dan/atau	11	Your *image has been* captured by *a video surveillance* installed in some parts of *our premise that are part of our security infrastructure*; and/or	**Grammatical error** - missing relative pronoun - SVA - wrong word order
Semasa transaksi perniagaan anda dengan kami; dan/atau	17	*Current business dealings* with us; and/or	**Grammatical error** - missing preposition
Secara amnya, melalui penggunaan cookies di laman web kami	18	Generally, through the use of cookies *of this site*	**Grammatical error** - wrong preposition

Table 4.7: Errors made by Google Translate in Text 4 (Notice) (continued)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
3. Tujuan pengumpulan, pemprosesan dan penggunaan Maklumat Peribadi anda.	19	3. The *purpose of the collection, processing and use* of your Personal Information.	**Grammatical error** - wrong word form - redundant article
Untuk memproses permohonan, pertanyaan dan permintaan *anda*;	20	For the processing of applications, inquiries and requests;	**Grammatical error** - missing possessive pronoun
Untuk rekod pengkalan data, penyelenggaraan dan kemaskini;	22	For *records database*, and *updates penyelenggaraan*;	**Lexical error** -different word choice (Malay word is not translated) **Grammatical error** - wrong word order
Untuk tujuan *pentadbiran* termasuk audit-audit, pemantauan penipuan dan pencegahan jenayah;	23	For *admin*, including audits, fraud monitoring and prevention of *criminals;*	**Lexical error** -wrong word choice
Untuk memastikan *sekuriti* dan keselamatan premis (*berkenaan* imej-imej CCTV);	26	To ensure *sekuriti* and security of the premises (*relevant* to CCTV images);	**Lexical error** -different word choice (Malay word is not translated) - wrong word choice

Table 4.7: Errors made by Google Translate in Text 4 (Notice) (continued)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Gabungan syarikat-syarikat kami dan anak syarikat-syarikat kami di dalam atau di luar Malaysia;	32	*Joined our companies* and our subsidiary companies within or outside Malaysia;	**Grammatical error** - wrong word order **Semantic error** -error in meaning
Penasihat, juruaudit, perunding, kontraktor, *pembekal produk dan perkhidmatan* kami sekadar makluman yang diperlukan sahaja;	34	*Advisor, auditor,* consultants, contractors, *suppliers and products* and services *necessary information just once;*	**Grammatical error** - wrong word form **Semantic error** -error in meaning
Mana-mana pemegang serah hak atau bakal penerima pindah milik atau pemeroleh syarikat atau perniagaan kami atau berkenaan dengan perlaksanaan korporat.	36	Any prospective assignees or transferees *or acquirer company* or our business or in connection with the implementation of corporate strategies.	**Grammatical error** - wrong word order
Polisi Privasi kami *telah menerangkan dengan lebih lanjut* mengenai bagaimana kami mengumpul, menggunakan dan memproses maklumat peribadi anda.	38	Our Privacy Policy *explains more about* how we collect, use and process your personal information.	**Grammatical error** - tenses **Semantic error** - error in meaning

Table 4.7: Errors made by Google Translate in Text 4 (Notice) (continued)

Source Text (BM) **(Notice)**	**Unit of translation (Vinay & Darbelnet, 1958)**	**Target Text- Google Translate (English)**	**Errors made by Google Translate**
Sila *lihat* Polisi Privasi kami di www.klsogo.com.my/privacy ("laman sesawang kami").	39	Please *see* our Privacy Policy www.klsogo.com.my/privacy ("our website").	**Lexical error** - different word choice **Grammatical error** - missing preposition
6. Akses, pembetulan atau menghadkan proses-proses Maklumat Peribadi anda	42	6. *Access, correction or limiting processes* your Personal Information	**Grammatical error** - wrong word order
6.1 Pada bila-bila masa anda boleh mengemukakan permintaan untuk akses, membuat pembetulan, mengemaskini atau menghadkan proses-proses maklumat peribadi anda dengan memaklumkan kepada kami secara bertulis kepada butir-butir hubungan seperti tertera di bawah:	43	6.1 At any time you can submit a request *for access, correction, update or limiting processes your personal information* by notifying us in writing to the contact details as shown below:	**Grammatical error** - gerund - wrong word order
Meluluskan atau menyambung keahlian anda dengan S Card;	47	*Adopt* or continue your membership *by* S Card;	**Lexical error** - wrong word choice **Grammatical error** - wrong preposition

Table 4.7 above showed that grammatical errors were identified as the most frequent errors made by Google Translate with twenty-five (25) errors, while there were nine (9) errors for lexical category and five (5) for semantic errors.

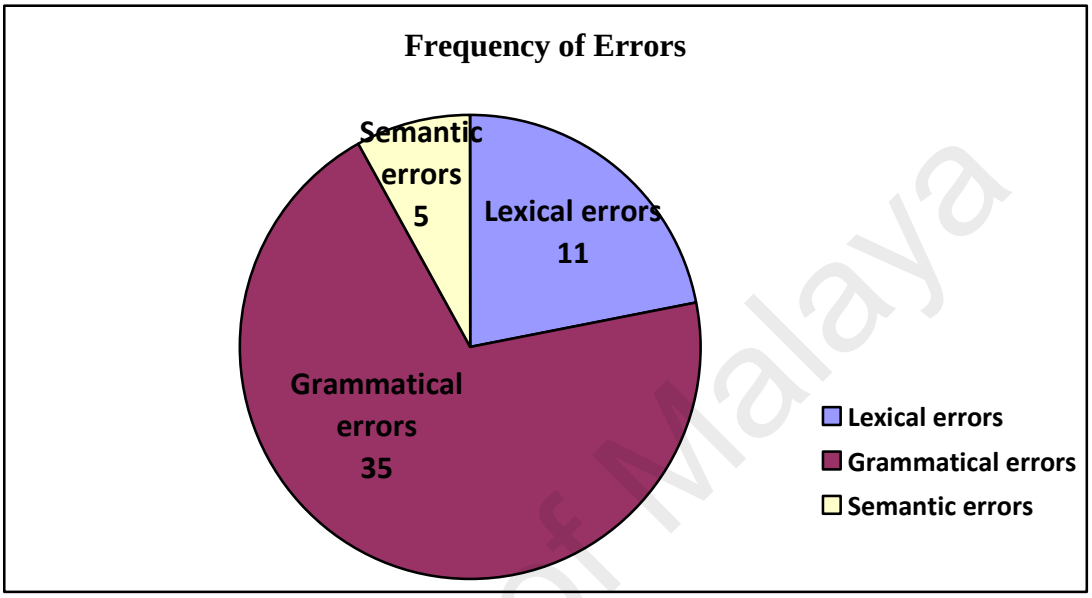

Figure 4.4: Frequency of Errors in Text 4 (Notice)

Further analysis on the specific identification of these three errors showed that wrong word order was shown as the most outstanding error made by Google Translate which occurred eight (8) times. Wrong word choices (lexical errors) also contributed to a high frequency of errors that was six (6). The other errors that were identified in the Google Translate output were wrong word form: three (3), the use of prepositions: four (4), different word choices: three (3), errors in meaning: three (3), gerunds: two (2) and the use of pronouns: three (3). Other errors that were rated with one (1) error for each such as subject-verb-agreement (SVA), redundant use of article and the tense used.

Table 4.8: Types of Error in Text 4 (Notice)

Errors Category	Types of Errors	Frequency
Lexical Errors	Wrong word choices	6
	Different word choices	3
Grammatical Errors	Gerund	2
	Wrong word order	8
	Subject-verb-agreement (SVA)	1
	Missing possessive pronoun	1
	Missing relative pronoun	2
	Missing prepositions	2
	Wrong prepositions	2
	Wrong word form	3
	Redundant article	1
	Tenses	1
Semantic Errors	Collocation	2
	Errors in meaning	3

4.3 Analysis of the Four Texts: To identify typical errors made by Google Translate when translating vocative and informative texts from BM into English.

A revised taxonomy of errors adopted from Keshavarz (1999) and Vilar et al. (2006) was used to identify the errors made by Google Translate. As shown in Figure 4.5 below, the most common type of errors produced by Google Translate for both notices and pamphlets was grammatical errors. A total number of frequencies for the four texts were sixty-one (61), where Text 4 had the highest frequency of errors which was twenty-two (22) errors as shown in Table 4.8.

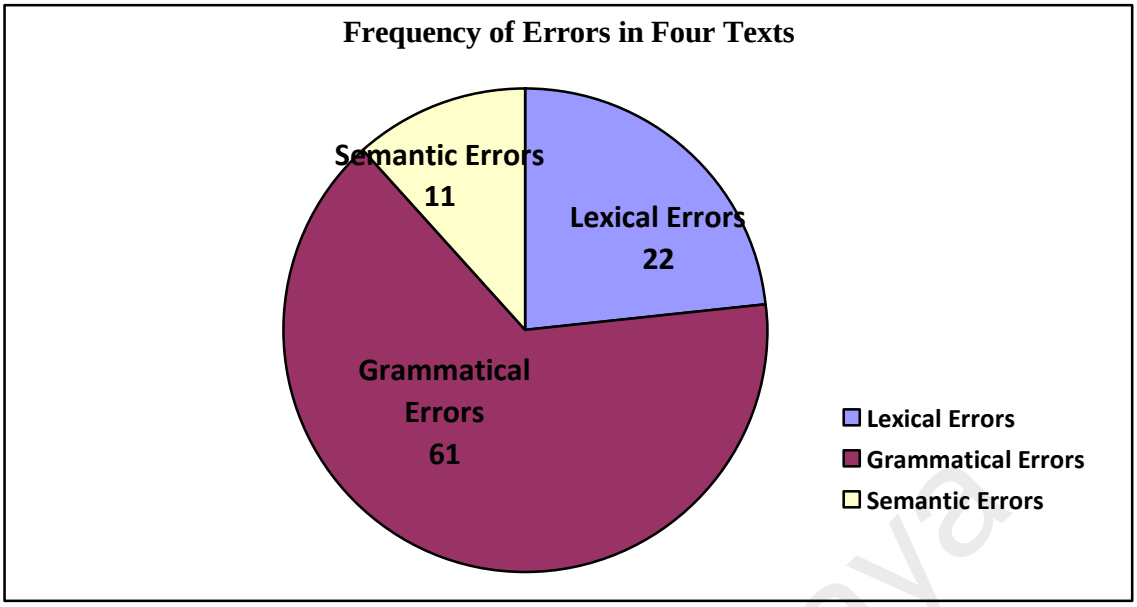

Figure 4.5: Total Frequency of Errors in Four Texts

Table 4.9: Total of Error Categories and Types of Errors for the Four Texts

Errors Category	Types of Errors	Text 1	Text 2	Text 3	Text 4	Total
Lexical Errors (22)	Wrong word choices	1	3	7	6	17
	Different word choices	-	-	2	3	5
Grammatical Errors (61)	Wrong word order	1	6	4	8	19
	Subject-verb-agreement (SVA)	2	2	3	1	8
	Missing prepositions	1	1	4	2	8
	Wrong prepositions	-	-	-	2	2
	Missing possessive marker	1	-	-	-	1
	Missing relative pronoun	1	-	2	2	5
	Missing object pronoun	-	1	-	-	1
	Wrong article	2	-	1	-	3
	Missing article	-	1	2	-	3
	Redundant article	-	-	-	1	1
	Gerund	-	1	-	2	3
	Tenses	1	-	1	1	3
	Wrong word form	-	1	-	3	4
Semantic Errors (11)	Errors in meaning	2	2	1	3	8
	Collocation	-	1	-	2	3
Total		12	19	27	36	94

The finding seems to be relevant with the findings of Eftekhar and Nouraey's (2013) study in which they recognised the most common error was grammatical errors when translating Iranian products label from English. The findings for this study also revealed that notices (Text 3 and Text 4) produced higher errors for the three categories of errors which was sixty-three errors in total compared to thirty-one errors in pamphlets (Text 1 and Text 2). This was not surprising as notices had longer sentences than pamphlets, hence the tendency for making errors was higher in the former. The finding of this study is also supported by Rensburg et al.'s (2012) study where they discovered that translations of newspaper articles and official letters produced more errors compared to PowerPoint slides or minutes of meetings as the former had longer sentences than the latter.

The grammatical errors were then divided into its sub-divisions, where it showed that there were nineteen (19) errors in wrong word order, eight (8) errors in subject-verb-agreement (SVA), three (3) errors related to gerund and tenses, as well as four (4) errors for wrong word form. The most highly rated of errors for wrong word order was found in Text 4 (Notice). Besides that, prepositions were found to have a similar frequency with SVA: eight (8) errors. It was then further divided into missing prepositions: eight (8) errors and two (2) for wrong prepositions. Word references had eight (8) errors in total where they were broken down to missing possessive marker: one (1) error, missing object pronoun: one (1) error and missing relative pronoun which had six (6) errors. Additionally, missing article and wrong article shared the same frequency which was three (3) errors and redundant article was reported to have one (1) error. Errors in tenses were made three (3) times in total.

As for the lexical errors, there were two sub-categorisations; wrong word choices with seventeen (17) errors and different word choices with five (5) errors in total for the four texts. The most incorrect use of word choices was from Text 3 (Notice)

with seven (7) errors, where some of the words translated use wrong words and were not relevant in the context of the source text. Some of the incorrect lexical items were also used repeatedly in the sentences and thus, led to the increased number of errors. In Text 3 (Notice), where *"Amanah Saham Nasional Berhad (ASNB)"* was translated as *"funds"* by Google Translate and occasionally used in other sentences led to the increasing number of errors.

In contrast, semantic errors were found in all texts respectively with a total number of eleven (11) errors. Some of these errors were words that Google Translate was unable to translate at all, hence, the Malay words remained untranslated. For example, in Text 4 (Notice), *"Untuk rekod pengkalan data,* ***penyelenggaraan*** *dan kemaskini"* was translated as *"For records database, and updates* ***penyelenggaraan****".* Google Translate was unable to identify the equivalent word to the source text in its dictionary and thus, remained unchanged and untranslated. There was also an interesting finding related to collocation errors which occured three times in total. For instance, the use of collocation in Text 4 (Notice), ***"seperti mata dan token"*** was translated as *"****as the eyes and tokens"****.* It was translated wrongly by Google Translate and directly affected the meaning of that particular sentence, even though the overall meaning of the rest of the text was acceptable.

To summarise, from the four texts, grammatical errors were found to have the highest number of errors with sixty-one errors (61), followed by lexical errors with twenty-two (22) errors and semantic errors with eleven (11) errors in total (refer to Figure 4.5). The common errors made by Google Translate were in terms of wrong word order: twenty one (21), wrong word choices: eighteen (18) errors, subject-verb-agreement (SVA) and prepositions: ten (10) errors, collocation: three (3) errors as well as errors in meaning with eight (8) errors. These were the most prevalent

errors that can be noted from Google Translate outputs when translating the vocative and informative texts.

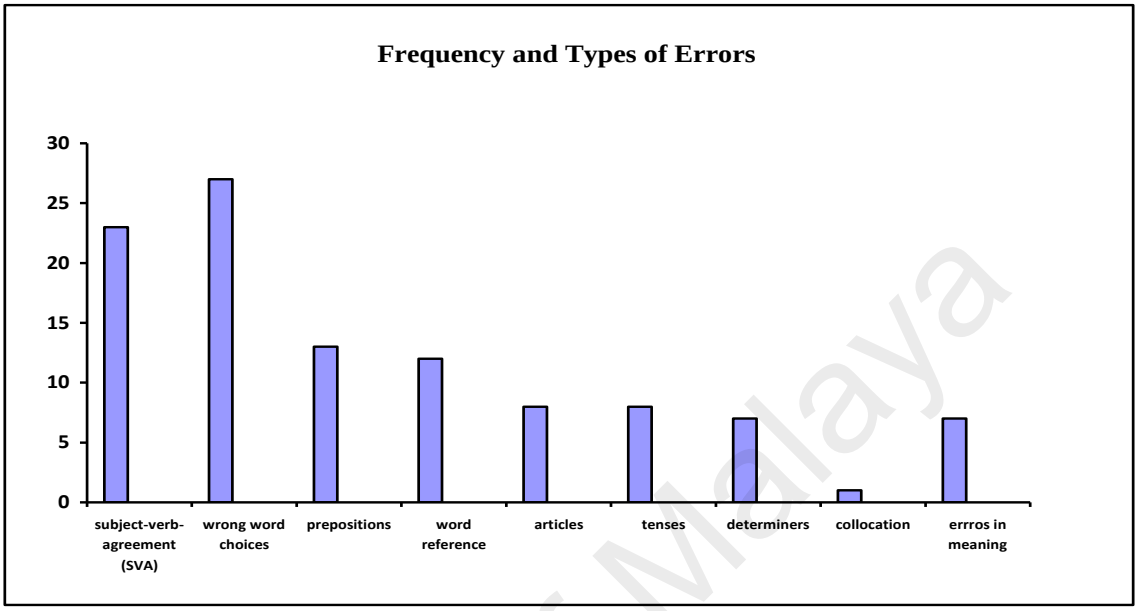

Figure 4.6: Types and Frequency of Errors

Thus, as shown in Figures 4.5 and 4.6, the analysis of grammatical errors had provided us an insight into the problems that Google Translate encountered when dealing with the data samples used in this study. Largely, wrong word order, subject-verb-agreement (SVA) and wrong word choice use were challenging for Google Translate to handle in the English output texts.

4.4 Rectification of the Errors (Human Translators post-editing)

As discussed in sections 4.2 and 4.3, the errors produced by machine translation, specifically Google Translate, for this study referred to the inappropriate translations at the lexical, grammatical (syntactic) and semantic levels. Since Google

Translate was unable to provide a satisfactory output for some instances, human editing was required in order to improve the quality of the target text. Sager (1994) asserts that the evaluation and revision of the machine translation output is either to eliminate the errors that affect the source text output or to make the output sound natural to the source text. Similarly, Melby (1987) says that post-editing is to revise the output of the text after the translating process. To achieve the second research objective, which was analysing the rectification made by the human translators, the outputs of the target texts produced by Google Translate were given to three human translators. All the three human translators are well-versed in translation and are expected to reproduce a satisfactory end product of the source text. Based on Keshavarz's (1999) and Vilar et al.'s (2006) categorisation of errors, the output of the Google Translate and human translators were compared. This was to examine how human translators rectified the errors made by Google Translate in order to improve the quality of the translation. The editing of the four texts by the three human translators was shown in detail in Appendixes A to E.

Table 4.10: Total of Error Categories and Types of Errors for the Four Texts

Errors Category	**Types of Errors**	**Total**
Lexical Errors	Wrong word choices	18
	Different word choices	5
Grammatical Errors	Word order	21
	Subject-verb-agreement (SVA)	10
	Prepositions	10
	Word references	9
	Articles	7
	Tenses	3
	Gerund	3
Semantic Errors	Errors in meaning	8
	Collocation	3

Table 4.10 above showed a total of number of errors for the three categories. It was further elaborated in the next sections 4.4.1 until 4.4.3.2 on its rectification made by the three human translators.

4.4.1 Grammatical Errors (Syntactic)

4.4.1.1 Prepositions

From the four translated texts provided, preposition use was the most frequent error picked up by the three human translators (HT). In unit translation 9 (Text 1), *'senak atau kesukaran untuk menelan'* was translated as *'indigestion or difficulty _ swallowing'*. In this target text produced by Google Translate, the preposition 'in' was omitted. The preposition 'in' was needed for this context to denote a motion that showed an action that came before. Hence, a preposition was needed to connect the adjective (difficulty) and noun (swallowing).

Another example, *'suara menjadi serak atau batuk melebihi 2 minggu terutama pada mereka yang tidak merokok'* was translated as *'voice become hoarse or cough more than 2 weeks, especially* ***in*** *those who do not smoke'.* For this output, the correct preposition needed to show the relation between 'mereka'/ 'those' (noun). So, when it was translated, the use of preposition was wrong because the word followed after was 'those' (demonstrative pronoun), and thus, a suitable preposition to be used was 'for' or 'among' (Longman Dictionary, p. 626). Referring to the source text, the three human translators rectified the errors as '...***(for)*** *non smokers/* ***(for)*** *those who do not smoke'*. Hence, the preposition 'for' was more accurate because it referred to intended group of people who received the consequences from other people's action.

Similarly, in unit of translation 3 (Text 3), *'Pihak kami mengucapkan terima kasih kerana menjadi pelanggan setia Amanah Saham Nasional Berhad ("ASNB")'* was translated as *'The wishes we thank you for being a loyal customer_ Amanah Saham Nasional Berhad ("funds")'*. The translated version of Google Translate mimicked the source text, but the use of preposition was missing in order to show the relation between 'pelanggan setia'/ 'loyal customer' (noun) and 'Amanah Saham Nasional Berhad (ASNB)' (proper/specific noun). Again, the simple preposition 'of' was needed and rectified by the three human translators in order to show the relationship of customers that referred to what or which entity (ASNB).

In units of translation 7,14 and 29 (Texts 3 and 4), *'mengumpul, menyimpan, memproses dan menggunakan Data Peribadi anda selaras dengan Polisi Privasi ASNB seperti yang dinyatakan dalam laman kami di www.asnb.com.my'* was translated as *'collect, store, process and use of your Personal Data in accordance with the Privacy Policy funds as set out in our website _ www.asnb.com.my'* and *'Atau anda boleh menghantar e-mel ke asnbcare@pnb.com.my'* was translated as *'Or, you can email _ asnbcare@pnb.com.my'*. Here, it can be seen that Google Translate was unable to determine or identify the preposition of place. Likewise, *'Sila lihat Polisi Privasi kami di www.klsogo.com.my/privacy ("laman sesawang")* (Text 4) was translated as *'Please see our Privacy Policy www.klsogo.com.my/privacy ("our website")'.* Therefore, it can be said that Google Translate was unable to translate the preposition of place for these particular units of translation and did not translate the preposition 'di' in (BM).

Another error of omission of preposition which was rectified and corrected by the three human translators was in unit translation 1 (Text 4). *'Notis Privasi ini adalah untuk memaklumkan anda mengenai polisi SOGO (K.L.) Department Store Sdn.Bhd (syarikat/kami) mengenai polisi untuk mengumpul, menggunakan, memproses dan mendedahkan Maklumat Peribadi anda'* translated as *'This privacy notice is to inform*

you about the policies _ SOGO (K.L.) Department Store Sdn. Bhd (a/us) on policies for collecting, using, processing and unwind disclose your Personal Information'. In this unit of translation, the preposition 'of' was omitted. This is because, Google Translate literally and directly translated from the source text. Hence the relationship between SOGO (K.L.) Department Store Sdn. Bhd (noun) and the policy was missing. Thus, this unit of translation was rectified by human translators as *'This Privacy Notice is to inform you regarding the policy of SOGO (K.L.) Department Store. Sdn. Bhd*.

Therefore, the incorrect use of prepositions or missing prepositions made by Google Translate was able to be recognised by the three human translators, and was rectified by them in the context of the source text. In BM, sometimes, the use of preposition was not emphasised as it was comprehensible and understood. Thus, it led Google Translate to translate it superficially, as what was stated in the source text. An example given in unit of translation 3, in Text 3, *'Pihak kami mengucapkan terima kasih kerana menjadi pelanggan setia Amanah Saham Nasional Berhad ("ASNB")* was translated by Google Translate as *'The wishes we thank you for being a loyal customer Amanah Saham Nasional Berhad ("funds")'*. Google Translate was unable to recognise that a preposition was needed in *'being a loyal customer Amanah Saham*' in the target text in order to produce a correct translation. With the help of human translators, the sentence was corrected as *'We would like to thank you for being a loyal customer* ***<u>of</u>*** *Amanah Saham Nasional Berhad ("ASNB").*

4.4.1.2 Articles

The second category of error that needed to be rectified by the three human translators was articles. In the unit of translation 15 (Text 1), *'untuk jenis kanser yang boleh dikesan awal, terdapat 80% peluang untuk hidup'* was translated as *'for this type of cancer can be detected early, there is* ***a 80%*** *chance of survival'*. The use of article 'a' was incorrect as it referred to a number which began with a vowel sound. Hence, it was rectified as *'...****an 80%*** *chance of survival'*.

In the same text, unit translation 20, *'mengekalkan berat badan unggul'* was translated as '*maintaining _ ideal body weight'*. Similar to unit translation 15, Google Translate was unable to identify the initial vowel sound in 'ideal' and literally translated the source text. Thus, all three human translators added the article 'an' correctly.

Besides, in unit translation 16 (Text 2), *'pusat ini dilengkapi dengan peralatan berteknologi tinggi yang setaraf dengan* ***hospital swasta*** *tetapi caj yang lebih murah dan sedikit atau tiada masa menunggu'* was translated as '*the centre is equipped with high-tech equipment equivalent to* ***the private hospital*** *but with charges cheaper and little or no wait time'.* The use of article 'the' was not really accurate because 'hospital swasta' (private hospital) was never mentioned earlier in other units of translation. Therefore, since it was not stated in other units of translation earlier, it was considered as a general term and the accurate article to be used was 'a' (Gaudart, H et al., 2006, p. 30). This is because, the readers or receivers of the target text did not know which one of the private hospitals was referred to as other private hospitals might not provide the service and thus, the use of 'the' should be replaced with 'a', to give a general sense to the noun (private hospital).

Additionally, in unit of translation 18 (Text 2), *'**Berkonsepkan 'rumah kedua'**, NCSM menyediakan tempat penginapan bagi pesakit kanak-kanak luar bandar dan penjaga mereka yang menjalani rawatan di Kuala Lumpur'* was translated as *'**Just as _ 'second home'**, NCSM provide accommodation for patients of rural children and their caregivers who are undergoing treatment in Kuala Lumpur'.* It can be observed that, Google Translate has directly translated the source text and a definite article was missing. Thus, it was rectified by the three human translators as 'with the concept of 'second home', 'using the 'second home' concept' and 'based on the 'second home' concept'.

Another rectification made by the three human translators was in unit translation 16 (Text 3). "*Sekiranya kami tidak menerima sebarang maklum balas daripada anda berhubung dengan penggunaan Data Peribadi anda dalam tempoh 14 hari dari tarikh notis ini, anda dianggap bersetuju dengan terma dan syarat yang dinyatakan di sini"* was translated as "*If we do not receive any feedback from you regarding the use of your Personal Data within 14 days of the date of this notice, you are deemed to agree **to _ terms and conditions** set forth herein"*. 'The' was used for things which has been mentioned previously. Google Translate omitted the definite article and it was rectified by human translators by adding the specific and definite article as in "*...have agreed to **the** terms and conditions…".*

It is seen that all the three human translators have no difficulty in rectifying the article errors made by Google Translate.

4.4.1.3 Word references

The use of references refers to a single word or group of words that has been mentioned earlier. Sometimes the use of references is to substitute similar words or ideas such as the use of pronouns. Pronouns refer to a noun, an individual, individuals, inanimate or animate things whose identity is made clear in the text (Gaudart, H et al., 2006, p. 17). There are a few types of pronouns which serve different functions such as demonstrative, object, relative, reflexive, and indefinite or possessive pronouns.

It can be seen that in unit of translation 11 (Text 1) the use of possessive marker was missing. The marker (apostrophe) was used to show ownership or possession as well as to inform about a sense of belonging. The example from the text, *"Batuk penghisap rokok yang bertambah teruk"* was translated as *"**Smokers** cough that worsens"*. In the example provided, apostrophe was needed to show that a thing or a belonging owns by whom as it usually answers the 'whose' question (Edwin, 2013, p.12). Thus, the three human translators were able to recognise the errors and corrected them by adding the possessive marker. Another example of omission of possessive pronoun can be identified in UT 9 (Text 3) as well as UT 20 (Text 4). The source text "*Adalah menjadi satu kewajiban mandatori untuk anda memberikan Data Peribadi kepada kami*" was translated as *"It shall be a mandatory obligation for you to provide_ Personal Data to us".* In this excerpt, Google Translate literally translated the source text and omitted the possessive pronoun (your) which later was rectified by the human translators.

Besides that, the most apparent error made by Google Translate was with the use of relative pronoun as in UT 15 (Text 1), UT 5 and UT 11 (Text 3) as well as UT 5 and UT 11 (Text 4). Relative pronoun is used to introduce a relative clause or describes a joining sentence (Edwin, 2013). In other words, it introduced a dependent clause that

provided more information of the independent clause such as a word, phrase or a sentence. It added an essential information about the clause or sentence that was modified and introduced in the main clause/ sentence. This information is important to understand the sentence's meaning correctly and thus, cannot be omitted in order to make sense of the sentence as in the source text. To illustrate, "*Untuk jenis kanser **yang** boleh dikesan awal…*" was translated as *"For this type of cancer _ can be detected early"*. It was similar to other units of translation especially in Texts 3 and 4 where it had the most number of missing relative pronouns. *"Justeru itu, kami ingin memaklumkan bahawa data peribadi dan segala maklumat anda **yang** telah diberikan kepada ASNB…"* was translated as *"Therefore, we would l**ike to inform you the personal data** and any information which you have given to the funds…"* as well as "*Anda berhak untuk mengakses Data Peribadi anda **yang** disimpan oleh kami…"* was translated as *"You have the right to access your Personal Data **held** by us…"*. Other example, "***Imej anda yang telah** dirakam oleh video pengawasan yang dipasang di beberapa bahagian premis kami…"* was translated as *"**Your image has been** captured by a video surveillance installed in some parts of our premise…"*. It showed that Google Translate was unable to identify the word '*yang*' in some of the source text and thus omitted it. '*Yang*' is used to explain the word or phrase before in order to stress on the importance of the sentence (Kamus Dewan Bahasa, 2013, p. 1811). Additionally, in unit translation 2 (Text 2), Google Translate omitted the object pronoun 'it'. *"Sila simpan risalah ini atau berikan kepada yang tersayang"* was translated as "*Please keep this brochure or **give to** a loved one*". It was later rectified by the three human translators correctly.

4.4.1.4 Tenses

Tense is a grammatical category which correlates with time or expresses certain actions in a time frame given. Certain actions are limited to a certain period of time, present or continual. The changes of tenses depend on the context given and changing the verb or action to a correct form is called conjugation (Lindner, 2005, p. 91). Based on the four texts, it can be seen that, Google Translate was unable to translate some of the units of translation according to their time usage such as in UT 7 (Text 1), UT 10 (Text 3) as well as UT 38 in Text 4. An example, *"Luka atau bisul yang tidak sembuh,* ***bertukar warna****, berdarah atau ulser"* was translated as *"Sores or blisters that do not heal,* ***changes color****, bleeds or ulcers".* In this unit of translation, the verb used indicated that the time or condition of verb was changed to past tense. Hence, the correct translation should be in the past form of "*changed color*" which was corrected by the human translators, except human translator 1 (HT1). Human translator 1 rectified it by changing the verb into a base form "*change color*".

Besides, in UT 38 (Text 4), "*Polisi Privasi kami* ***telah menerangkan*** *dengan lebih lanjut mengenai bagaimana kami mengumpul, menggunakan dan memproses maklumat peribadi anda"* was translated by Google Translate as *"Our Privacy Policy* ***explains*** *more about how we collect, use and process your personal information*". The source text was intended to say that the policy was already explained in detail previously, hence, a present perfect tense should be used. In contrast to the two human translators, human translator 2 (HT2) used a present form for this unit which was similar to the raw output.

Further analysis of the texts revealed conditional sentences. A conditional consists of two clauses and are divided into open conditional or counterfactual conditional, where the former is seen to be realistic possibility while the latter is seen to

be impossible and contrast to the fact (Trask, 1993, p. 55). In unit translation 10 (Text 3), *"**Jika anda enggan memberikan** dan tidak membenarkan Data Peribadi anda digunakan bagi tujuan-tujuan yang dinyatakan, ia boleh menghalang anda untuk melanggan atau menggunakan sepenuhnya produk dan perkhidmatan kami yang mana pihak kami tidak bertanggungjawab dan tidak boleh dipertanggungjawabkan **sekiranya** perkara tersebut berlaku*". The raw output produced by Google Translate was *"If you refuse to leave and do not allow use of your Personal Data for the purposes that have been disclosed, it may prevent you to subscribe to or use of our products and services which we are not responsible and cannot be held responsible **if the event occurred**".* Here, it was noted that, when an action seems to be realistic, a present form should be used. For the first clause of the sentence, Google Translate was able to retain a present form but for the second clause it used the past form and thus, was corrected by the three human translators. It showed that the human translators were aware of the tenses used for this unit of translation.

4.4.1.5 Gerund

Trask (1993, p. 118) defines gerund as an '*-ing*' form of a verb where it serves as a verbal noun. It retains its ability to take verbal arguments, adverbs or complements. There were some sentences that were wrongly translated by Google Translate when it dealt with gerund. It can be seen in UT 16 (Text 2) and UT 1 and UT 43 (Text 4). In unit of translation 16, *"...dengan caj yang lebih murah dan sedikit atau tiada masa menunggu"* was translated as *"...with chargers cheaper and little or **no wait time**"*. Other example was "*...untuk mengumpul, menggunakan, memproses dan mendedahkan Maklumat Peribadi anda"* was translated as *"**for** collecting, using, processing and **unwind disclose** your Personal information".* Google Translate was unable to identify

the use of gerunds and literally translated the sentence but, the three human translators were able to correct the errors.

4.4.1.6 Subject-verb-agreement (SVA)

The incorrect use of subject-verb-agreement (SVA) made by Google Translate can be seen in all of the four texts. Google Translate was unable to identify the subject that agreed to the verb and thus translated it as a singular or plural form. The pertinent examples were taken from UT 10 and UT 11 (Text 1), UT 6 and UT 18 (Text 2), UT 4, UT 5 and UT 6 (Text 3) as well as UT 11 (Text 4). *"Suara menjadi serak*" was translated as *"**Voice become** hoarse"*. Human translators 1 and 2 (HT1 & HT2) translated the sentence as a passive voice while human translator 3 (HT3) corrected the sentence followed the subject which referred to the 'voice' as a singular. Other examples were, "*Berkonsepkan 'rumah kedua', NCSM menyediakan*" was translated as *"Just as 'second home' **NCSM provide***", "*Sebagai pemegang unit ASNB, anda berakujanji, mengakui"* was translated as *"As the Holders of the funds, **you undertakes, acknowledges"*** as well as "*Diurus oleh **seorang kaunselor** berpengalaman*" was translated as *"Managed by **an experienced counselors***". With the examples given, it seems that Google Translate was unable to identify the subject of the sentence, and thus produced it incorrectly. All the errors were corrected by the three human translators correctly.

4.4.1.7 Word order

Word order is a linear sequence of words in which it occurs in a constituent or in a sentence. It is usually related to the arrangement of the order of Subject, Verb and Object (Task, 1993, p. 306). It was notable that the common errors made by Google Translate when translating notices and pamphlets were in the form of word order arrangement. This included the use of conjunction. The incorrect translations produced by Google Translate were also due to direct translation or mimicked the structure of source text where some of the sentences had complex sentences. For example, in UT 4 (Text 1), *"Terdapat ketulan yang tumbuh secara tiba-tiba* ***dan tidak hilang diikuti dengan sakit****"*. Google Translate translated it as *"There are lumps that grow* ***suddenly disappear and not followed by pain****"*. Here, the main clause was "There are lumps that grow suddenly" and the subordinate clauses were "the lumps do not disappear" and "the lumps are painful". Due to the fact that the syntactical structure of the source language was employed by Google Translate, it affected the meaning too. As for the three human translators, they were able to justify the meaning of the sentence and corrected the output of Google Translate.

In Text 2, there were five errors related to wrong word order. In UT 4 and 17, *"Pusat Sumber dan Kesejahteraan"* was translated as *"Resource Centre and Welfare"* and *"Rumah Harapan Kanak-kanak"* was translated as *"Hope childhood home"*. The placement of conjunction was wrong and Google Translate mimicked the source text which resulted to wrong translation. Human translators corrected them as *"Resource and Welfare Centre"* as well as *"Children's Home of Hope"* from both human translators 2 and 3, while human translator 1 corrected it as *"Children's House of Hope"*. For UT 13 in Text 2, *"Pasukan pakar teknologi perubatan kami…tentang penularan (metastasis) kanser* ***di organ badan dan tulang"*** was translated as *"Our*

team of experts in medical technology…infection (metastasis) of cancer ***in the body organs and bones****"*. It showed that Google Translate translated it literally and it also discovered that all the three human translators were unable to rectify it correctly. The three of them edited it as "*in the organs and bones"* for human translators 1 and 2, while "*in the body and bones*" from human translator 3. The researcher believed that the corrected output was "*in the organs of the body".* Another example from Text 2 in the unit of translation 16, *"Pusat ini dilengkapi dengan peralatan berteknologi tinggi setaraf dengan hospital swasta tetapi dengan* ***caj yang lebih murah*** *dan tiada masa menunggu"* was translated as *"The center is equipped with high-tech equipment equivalent to the private hospital but with* ***charges cheaper*** *and little or no wait time".* Again, the wrong word order in this unit of translation due to the fact that Google Translate literally translated and mimicked the source text and human translators had no problem in correcting them.

Besides, in Text 3, the wrong word order can be seen in UT 4 and UT 12. *"Dengan berkuatkuasanya Akta Perlindungan Data Peribadi 2010 ("Akta"), ASNB sebagai pengguna data dikehendaki untuk memastikan privasi data peribadi pelanggan yang telah diberikan atau sedang digunapakai atau diselenggara oleh ASNB dipelihara dan dilindungi"* was translated by Google Translate as "*With the Personal Data Protection Act 2010 ("the act"), as* ***the funds required to ensure the user data privacy of personal data of customers who have been granted or are being adopted or maintained by the funds are preserved and protected****"*. It resulted to literal translation produced by Google Translate and thus, was corrected by the three human translators correctly.

In Text 4, it can be seen in UT 11, 22, 36, 42 and 43. Similar to the other examples provided, the wrong word order was due to mimicking the source text. *"Imej anda yang telah dirakam oleh video pengawasan yang dipasang…"* was translated as

"Your image has been captured by ***a video surveillance*** *installed…".* Besides that, "*Untuk rekod pengkalan data, penyelenggaraan dan kemaskini"* was translated as "*For r****ecords database, and updates*** *penyelenggaraan*". Another example from Text 4 was, *"Pada bila-bila masa…akses, membuat pembetulan, mengemaskini atau menghadkan proses-proses maklumat peribadi anda…"* was translated as "*At any time…access, correction, update or limiting* ***processes your personal information****…".* The errors were then rectified by the three human translators rightly.

4.4.1.8 Wrong word form

The wrong word form was found mostly in Text 4 (Notice), in units of translation 10, 19 and 34. Google Translate was unable to differentiate some parts of speech such as nouns, verbs or adjectives. In UT 10, "*…pertandingan, kajian, promosi, soal selidik atau komunikasi dengan kami…*" was translated as "*…contests, surveys, promotions, surveys, promotions, surveys or* ***communicate*** *with us…".* Similar to UT 15 (Text 2), *"Rawatan untuk pesakit-pesakit kanser termasuk radioterapi…"* was translated by Google Translate as *"Treatment for cancer patients* ***including****…".* From the two excerpts given, it showed that Google Translate was unable to differentiate between nouns and verbs.

4.4.2 Lexical Errors

This is when Google Translate produced the raw output of the source text by using wrong word choices or different word choices. The use of wrong word choices somehow affected the meaning of the text. On the other hand, different word choices are when Google Translate did not find the equivalent words as in the source text.

4.4.2.1 Wrong word choices

The examples of wrong word choices can be seen in UT 2 (Text 1), UT 4, UT 7, UT 10 (Text 2), UT 3, UT 4, UT 5, UT 6, UT 7, UT 10 and UT 11 for Text 3 while UT1, UT 7, UT 10, UT 23 and UT 26 (Text 4). As in Text 2, "*Perubahan dalam '**corak'** pembuagan air besar atau air kecil yang berterusan dalam jangkamasa ke 2-3 bulan*" was translated as *"Changes in the **'twist'** defecation or urination continued over a period of 2-3 months*". The word 'corak' in the source text refers to a pattern or a characteristic (Kamus Dewan, 2013, p. 288). However, it was mistranslated as 'twist' which provided a different meaning. Thus, the three human translators edited the unit of translation as '*Continuous changes in the **'pattern'** of defecation or urination over a period of 2-3 months*".

Another example that can be found in the unit of translation was the use of *'kesejahteraan'* that was "*Pusat Sumber dan Kesejahteraan*" or '*kelas kesejahteraan seperti yoga dan qigong*". It showed that Google Translate translated it as '*welfare*' which means a different thing. In Text 3, '*ASNB'* which was translated as *'funds'* was repeated in the text: UT 4, UT 5, UT 6 and UT 7. When the word was repeated, the tendency of Google Translate to produce the same word was high and thus, contributed to errors in word choices. However, Google Translate was able to translate the proper name of Amanah Saham Nasional Berhad but was unable to produce the correct abbreviation for ASNB, where '*ASNB'* was translated as '*fund'*.

Besides that, "*berkaitan dengan apa-apa perkhidmatan, transaksi, pertandingan, **kajian**, promosi, soal selidik atau komunikasi*" in UT10 (Text 4) was translated as *"in connection with any services, transactions, contests, **surveys**, promotions, **surveys,** or communicate*". There is an equivalent of word to '*kajian'* which is *'study or research'*, but failed to be recognised by Google Translate as well as the

three human translators. All of them were unable to correct the word where they rectified it as '*survey'.*

Moreover, in UT 23 and UT 26 (Text 4), *"Untuk tujuan pentadbiran termasuk audit-audit, pemantauan penipuan dan pencegahan jenayah"* was translated as *"For admin, including audits, fraud monitoring and prevention of* ***criminals****"*. "*Untuk memastikan sekuriti dan keselamatan premis (berkenaan imej-imej CCTV)"* was translated as *"To ensure sekuriti and security of the premises (****relevant*** *to CCTV images)".* As shown by the examples provided, the three human translators corrected the errors correctly. They rectified the error as "*crime"* and "*regarding / in relation to".*

4.4.2.2 Different word choices

Google Translate used different word or remained the source text word by not translating it when there was no equivalent of words in the source text. In UT 8 (Text 3), ***"menzahirkan*** *dan berkongsi Data Peribadi anda*" was translated as ***"demonstrate*** *and share your Personal Data".* The word '*menzahirkan*' in the context of the source text meant to give openly the information. Conversely, 'd*emonstrate*' was used to show something that can be seen and was not meant for something abstract such as data or information. Since Text 3 was a notice, a correct word to be used was '*disclose*' which can be found in any legal or formal document. All the three human translators were able to correct the error made. Also found in UT 39 (Text 4), *"Sila* ***lihat*** *Polisi Privasi kami*" to "*Please* ***see*** *our Privacy Policy*". It implied that the target text sentence was unsound and was literally translated by Google Translate. To note on the changes made by the three human translators, only human translator 3 (HT3) was said to have rectified the error made by changing the word to *"****Do visit*** *our website*" and it seems more accurate as compared to the other two human translators. Human translator 1 (HT1) changed to

"Please read" which was not quite accurate according to the source text context while human translator 2 (HT2) remained to use the same output.

In UT2 2 and UT 26 in Text 4, it showed that Google Translate was unable to identify the similar word as in the source text and chose to remain the word without translating it. For example, *"Untuk rekod pengkalan data,* ***penyelenggaraan*** *dan kemaskini"* and *"Untuk memastikan* ***sekuriti dan keselamatan*** *premis (berkenaan imej-imej CCTV)".* The former was translated as *"For records database, and updates* ***penyelenggaraan****"* and the latter as *"To ensure* ***sekuriti*** *and security of the premises (relevant CCTV images*)". The second instance showed that Google Translate remained the word use as in the source text due to the word used after was *'keselamatan'*. Hence, this created a challenge for Google Translate to make sense of the word given. Thus, the three human translators edited the error to '*mantainance*' for '*penyelenggaraan*' and '*security*' for '*sekuriti'*.

4.4.3 Semantic Errors

4.4.3.1 Errors in meaning

The semantic errors happened when Google Translate mistranslated the sentence which resulted in an incorrect use of lexical items or wrongly translated the sentence structure. A unit of translation then can have a combination of the errors in terms of lexis, syntactic and semantic. An example of semantic meaning due to incorrect translation of the word structure was in UT 4 (Text 1) where the mistranslation produced by Google Translate changed the meaning of the source text. Another example that can be analysed was in UT 4 (Text 3). "*Dengan* ***berkuatkuasanya*** *Akta Perlindungan Data Peribadi 2010* ***("Akta"), ASNB*** *sebagai pengguna data dikehendaki untuk memastikan* ***privasi data peribadi pelanggan*** *yang telah diberikan atau sedang*

digunapakai *atau diselenggara oleh ASNB dipelihara dan dilindungi.* The raw output produced by the machine was "*With the Personal Data Protection Act 2010 ("**the act")**, as the **funds** required to ensure the **user data privacy** of personal data of customers who have been granted or are **being adopted** or maintained by the funds are preserved and protected".* For this example, firstly, Google Translate did not translate the word *'berkuatkuasa'* and omitted it. Secondly, the translation of '*Akta*' was registered as a common noun and was not as a proper noun by Google, although it was translated correctly. Again, the *ASNB* was translated as *'funds'* and the translation of *'pengguna data'* was translated as *'user data'* that gave a whole different meaning for that particular unit of translation. Hence, the three human translators had rectified or reproduced the output of Google Translation with a more accurate sentence.

4.4.3.2 Collocations

The use of collocation also resulted in incorrect translation produced by Google Translate which affected the meaning of the texts too. In UT 7 (Text 4), "*Maklumat peribadi yang berkaitan dengan S Card **seperti mata dan token** yang dikumpul atau ditebus oleh anda"* was translated as *"Personal information related to the S Card **as the eyes and tokens** collected or redeemed by you"*. Collocation is idiosyncratic or specific language (Trask, 1993, p. 49) that comes or collocates together which provides a specific meaning. In the source text, 'mata' referred to points obtained by customers when they used their membership card for purchasing and will benefit them. However, 'mata' was translated literally and thus, changed the meaning of the whole sentence. It showed that Google Translate was unable to translate a collocation which was later corrected by the human translators. It was also found in UT 10 (Text 2) where *"...pemeriksan payudara sendiri dan cara pemakaian payudara palsu..."* was

translated as *"breast self-examination and how to use **fake breast**..."*. Google Translate literally translated it and for this particular unit of translation too, the three human translators were unable to produce the correct collocation in the context of breast cancer. Human translators 1 and 3 maintained using 'fake breasts' while human translator 2 used 'artificial breasts'. The researcher did a research on it and found that it should be 'breasts prosthesis' in the context of breast cancer.

4.5 Translation Quality Assessment (Human Translators)

The raw output of the target texts produced by Google Translate together with the source texts were given to the three human translators. They were asked to rectify any necessary mistranslated output produced by Google Translate based on units of translation (UT). As mentioned in section 2.4, UT was used to segment the sentences for easier identification or errors, but, the chain of the UTs was related to one another. The three human translators are well-versed in translating and translation, as they have experience in translating works. Cases where the three human translators differ over on using certain word choice or grammar issue should be considered. This is because, the difference in word choices is still acceptable if the output is relevant to the source text.

Keshavarz (1999) says, error analysis is collecting samples, identifying the errors and classifying the errors into categories and finally evaluating the errors (which means edited by human translations) where necessary. Following the Keshavarz's and Vilar et al.'s framework, the outputs of Google Translate were given to the three human translators. However, the three human translators may have different approaches in correcting the errors or units of translation. To rank the three human translators, Human Translator 2 and 3 (HT2 & HT3) seem to produce better and accurate translation, despite the fact that the three of them did rectify some of the errors incorrectly. The

clearest example from Human translator 1 (HT1) was when dealing with the use of determiner. 'Any' should be used for a singular form. It is similar to 'every'. However, HT1 failed to identify the error. Besides, HT1 also failed to identify the changes of tense as in UT 7 (Text 2). Basically, HT1 did not really rectify the errors on grammatical aspects such as determiner and tenses.

In terms of lexical level, HT1 had inappropriately edited the words to cater the meaning of the source text. For example, *"Berkonsepkan 'rumah kedua', NCSM menyediakan tempat penginapan bagi pesakit kanak-kanak luar Bandar dan penjaga mereka yang menjalani rawatan di Kuala Lumpur".*

HT1: *With the concept of 'second home', NCSM provides* ***accommodations*** *for* ***outskirt*** *child patients and their guardians who are undergoing treatments in Kuala Lumpur*.

HT2: *Using the 'second home' concept, NCSM provides accommodation for children patients from rural areas and their caregivers who undergo treatment in Kuala Lumpur*.

HT3: *Based on the 'second home' concept, NCSM provides accommodation for patients of rural children and their caregivers who are undergoing treatment in Kuala Lumpur*.

From the excerpts given, the researcher believed that the lexis used by HT1 did not correct, although it was a synonym and it was not relevant to this context as it sounded awkward.

Google Translate is a statistical machine translation where it translates the output literally. In UT 3 (Text 2), *"Pusat Kebangsaan Kanser Malaysia (NCSM)*" was translated as "*National Cancer Malaysia* (NCSM) and in UT 3 (Text 3), "*Amanah Saham Nasional Berhad ("ASNB")* was translated as "Amanah Saham Nasional Berhad *("funds").* For the former example, the output produced by Google Translate was

correct, but the registered name used was “National Cancer Society of Malaysia”. The latter, on the other hand, inexplicably changed the abbreviation into an unrelated word. The three human translators acknowledged the error and thus, provided the correct translation.

In contrast, the three human translators had provided minimal attention to semantic analysis. Since this study used notices and pamphlets, the language utilised was more formal. It was identified from the findings that there was no figurative language used, except the use of collocations in UT 10 (Text 2) and UT 7 (Text 4). The collocation in Text 2 was unable to recognise by the three human translators where they remained the similar word as produced by Google Translate but, was able to rectify collocation in other units of translation.

Collina (2009) believes that translators should have knowledge on the function of the source text and the intended audience that received the information. The criteria that need to be considered by human translators are the target-language user, the function and adequacy of the text and ability to transfer a non-specific or specific content. House (1981, p. 71) stressed on the equivalence of both texts (source text and target text) where she explains that equivalence should look into semantic and textual aspect as well as syntactic and lexical aspects. They cannot be viewed alone as each language has different linguistic items and sometimes ambiguous in use. This is because, both source and target texts should match with one another in function as in each text itself had its particular function such as expressive, informative or vocative. Relating this to the quality of end products produced by the three human translators on rectification of the errors made by Google Translate, human translator 2 and 3 (HT2 and HT3) had the most acceptable and accurate translations as compared to human translation 1 (HT1). In fact, based on the simple questionnaire given to the three human translators, HT1 agreed that the meaning of the target texts (English) was preserved or

maintained as in the source text, and little editing was needed. She further explained that the message conveyed by using Google Translate was maintained although some terms in the target texts did not necessarily correct based on the genre on the texts themselves.

In comparison, human translator 2 (HT2) and human translator 3 (HT3), strongly disagreed that the meaning of the target text was completely preserved as in the source text. HT2 believed that among the four texts, Text 4 was the text that needed the most rectification. Likewise, according to HT3, all outputs produced by Google Translate especially Text 4, needed correction on its lexical, morphological syntactical and semantic aspects. It was true because Text 4 had longer sentences or units of translation and thus, produced a high frequency of errors (see Figure 4.4).

Therefore, human translator 3 (HT3) had rectified the errors made by Google Translate accurately as in the source text and followed by human translator 2 (HT2) based on the findings provided. Conversely, human translator 1 (HT1) produced satisfactory output as there were still units of translation which needed correction. After all, all the end outputs which were rectified by the three human translators were acceptable based on the source texts given. However, human translation 1 (HT1) needed to pay more attention to the grammatical aspect of a text.

4.6 Conclusion

This chapter has identified that there were three errors categories: lexical, grammatical and semantic. All of the error categories were found in both pamphlets and notices. The results revealed that grammatical errors hold the highest frequency. The longer the text was, the higher the tendency of making errors by Google Translate. The raw outputs of the target texts were rectified by the three human translators and it showed that some of the errors were not corrected.

CHAPTER 5

CONCLUSION

5.0 INTRODUCTION

This chapter presents an overview of the study, the implications as well as the suggestions provided for further research.

5.1 Overview of the Findings

Translation is rendering the meaning of a source text into another target text in which is acceptable by the readers in a way that the author or translator intended the text (Newmark, 1988). Similarly, Nida and Taber (1969) state that translation is when a target text is transferred or transformed into the surface structure of the target text that involves three stages (analysis, transfer and re-struture). The use of machine translation (Google Translate) is an alternative way for most students or language users who need to translate their works from one language to another language without paying for any cost. This is because, as mentioned earlier by Nurul Hidayah, a representative from ITBM, the cost of translation is a bit pricey and it depends on the length and type of text. Hence, the best choice is using Google Translate which is available freely online. However, the question arises as to what extent does Google Translate help in producing a good output from a source text to a target text. Given the findings from all the texts used in this study, it can be said that Google Translate was able to produce a good translation output especially when the texts use simple sentence structures. As attested by human translator 1, using Google Translate provides a general idea of the text as

well as choices of words for different word categories such as noun, adjective, verb or adverb, making it easier especially for translators to translate according to the context.

Based on the questionnaire given to the three human translators, they agreed that the use of Google Translate helps people to produce output of the target text despite the fact that the result did not make any sense sometimes because Google Translate did not understand the context of the source text. From an example given by human translator 2 in the questionnaire , *'Fire! Fire!*' was translated to *'Tembak! Tembak!'* and not *'Api! Api!*' or *'Perubahan dalam '**corak**'* into '*Changes in the '**twist**'* (UT 2, Text 1). Here lies the weakness of machine translation where it connected two languages word by word but not meaning by meaning, particularly in text that contained figurative elements such as collocations as in '*mata dan token*' becomes '*eyes and token*' in UT 7 of Text 4.

Using the error analysis framework proposed by Keshavarz (1999) and Vilar et al. (2006), the samples of the errors of the four texts made by Google Translate were examined. The errors were categorised into lexical, grammatical and semantic aspects and each of the errors were examined in detail. Based on the findings, it was revealed that Text 1 and Text 2, an informative text (pamphlets) produced fewer errors in terms of grammatical aspects as compared to Text 3 and Text 4: vocative (notices). Both text types serve similar functions but Text 1 and Text 2 focused more on providing information. Additionally, pamphlets were shorter than notices, hence, the number of errors produced by machine translation for informative text was lower than vocative texts (Chun-ling, S. 2007 & Chen et al., 2016). It is also found that the most significant error made by Google Translate was in terms of wrong word order. In some units of translation (in pamphlets and notices), Google Translate was unable to recognise the basic word order of the source text (BM) due to the order of elements that appeared to be completely free or are used for different grammatical purposes as well as the subject

cannot be identified by Google Translate (Trask, 1993, p. 29). For instance in unit translation 4 (Text 1), *"Terdapat ketulan yang tumbuh secara tiba-tiba* ***dan tidak hilang diikuti dengan sakit****"* was translated as *"There are lumps that* ***grow suddenly disappear and not followed by pain****"*. This is a very interesting error made by Google Translate in failing to recognise that the negative *'tidak'* (not) referred to the verb *'hilang'* (appear), instead it erroneously attached it to the final verb phrase in the source text "*diikuti dengan sakit",* resulting in a totally incorrect translation of the characteristic of a lump which brought pain to the sufferer. Thus, literally translated or mirrored the source texts were resulted in the wrong structure of the target text output which affected the meaning as well.

There were also errors that were overlooked by the three human translators and the researcher believed that the errors were deemed to be rectified. For instance, in the context of breast cancer, the medical term uses for *'cara pemakaian payudara palsu'* in UT 10 (Text 2) should be translated as *'how to use breast prosthesis'*. The researcher has referred to the official website related to breast cancer and found that it was the correct term or collocation to be used. In the same text, UT 13, it was found that the three human translators did not produce a correct phrase. '*kanser di organ badan dan tulang'* should be translated as '*cancer in the organ of the body'.* Besides, in UT 10 (Text 4), '*kajian*' was translated as 'survey' by the three human translators. There is an equivalent of word to *'kajian'* such as 'research' and the researcher considered that as a corrected term which was deemed to be rectified. Nonetheless, all the three human translators were aware on the specific and registered name uses for Persatuan Kebangsaan Kanser Malaysia (NCSM). Although in the pamphlet (Text 2, UT 3) the source text used 'Pusat Kebangsaan Kanser Malaysia' which translated as "National Cancer Centre Malaysia (NCSM)", the three human translators were able to identify it and corrected it correctly.

This study also presents the result of post-editing task where three human translators rectified or edited the raw outputs of Google Translate with access to the source text. The error rectifications made by the human translators was made based on the error categorisation's taxonomy. The results showed that human translator 1 (HT1) produced a less accurate output as compared to the other two human translators. The reason is that HT1 did not correct some of the errors found in Google output which deemed to be corrected for a better output. Among the three human translators, human translator 3 (HT3) produced the best edited translation.

5.2 Implications of the Study

This study focuses on identifying errors made by Google Translate in vocative and informative texts (notices and pamphlets) and how human translators rectify the errors. Based on the answers from the questionnaire provided, they believed that any text that was translated by machine translation needed to be edited by human. This is to ensure the target text is accurate from the smallest unit of meaning to the whole purpose of the text.

In fact, using Google Translate or other machine translations help to save time as it produces a quick translation of the text as well as it provides synonymous words for the output. The disadvantage of this tool is the inaccuracy of language use where the structure of the output sometimes mirrors the source text and literal translation is often produced.

Since human editing is still important for a better quality of the target text, the knowledge brought by the translators also plays an important role. The translators must be well-versed in both languages (source and target text). For this study, the human translators were chosen on the basis of their expertise works in translation.Their

intuition and style of translating differ from each other and cannot be judged on that matter. Hence, it is expected that they use different approaches in translation but maintained closely the original meaning of the source text.

5.3 Recommendations for Future Research

For Malay to English translations, Google Translate has difficulties with certain types of errors as shown in this study. It would be interesting to test other machine translations to compare the rate of accuracy to find out which of them is most able to produce translation outputs with the least errors. Further investigation also could explore other genres of text types and use a larger volume of texts.

5.4 Conclusion

In short, based on the findings and human translators' perceptions on using machine translation, it is clear that the use of Google Translate can aid or assist students or language practitioners to obtain a general idea of the text despite some of the sentences produced by Google Translate requires editing. The same patterns of errors identified by both from the research and human translators prove that these are the weaknesses of Google Translate when translating a text from BM into English. A practical outcome of this study is to make Google Translate aware of these problems and to urge them to improve their machine translations by identifying potential linguistic features and thus, produce more accurate translation outputs.

REFERENCES

Abdel Monem, A., Shaalan, K., Rafea, A. & Baraka, H. (2009). Generating Arabic text in multilingual speech-to-speech machine translation framework. *Journal of Machine Translation*, 22(4), 205-258.

Aiken, M., & Balan, S. (2001). An analysis of Google Translate accuracy. *Translation Journal*, l6(12). Retrieved from http://www.translationdirectory.com/articles/article2320.php

American Translation Association. (2010). Retrieved from www.atanet.org

Angela, C., Wang, L., Tiago L,. et.al. (2015). *A Linguistically Motivated Taxonomy for Machine Translation Error Analysis*. Retrieved from http://link.springer.com/article/10.1007%2Fs10590-015-9169-0

Aziz, W., Sousa, S., & Specia, L. (2012). PET: A Tool for Post-editing and Assessing Machine Translation. *International Conference on Language Resources and Evaluation*, 3982-3987. Turkey. Retrieved from http://www.mt-archive.info/Aslib-2012-Aziz

Balkan, L. (1998). Machine-aided Translation. In Baker, M. (Ed.), *Routledge Encyclopedia of Translation Studies*, 134-148. New York: Routledge.

Bell, R.T. (1991). *Translation and Translating*. London: Longman.

Bojar, O. (2011). Analyzing Error Types in English-Czech Machine Translation. *The Prague Bulletin of Mathematical Linguistics*, 95, 63-76. Retrieved from https://ufal.mff.cuni.cz/pbml/95/art-bojar.pdf

Brown, H.D. (1994). *Teaching by Principles an Interactive Approach to Language Pedagogy (2nd ed.).* New York: Pearson Education.

Brown, H.D. (2007). *Principles of Language Learning and Teaching (5th ed.)*. New York: Pearson Education.

Catford, J. C. (1965). *A Linguistic Theory of Translation: An Essay to Applied Linguistics*. London: Oxford University Press.

Chung-ling, S. (2007). Teaching Translation of Text Types with MT Error Analysis and Post-MT Editing. *Translation Journal*, 11(2). Retrieved from http://trnslationjournal

Collina, S. (2009). Translation Quality Evaluation: Empirical Evidence for a Functionalist Approach. *The Translator*, 14(1), 97-134. Retrieved from https://benjamins.com/#catalog/journals/target.21.2.02col/details

Condon, S., Parvaz, D., Aberdeen, J., et.al. (2010). *Evaluation of Machine Translation Errors in English and Iraqi Arabic*, 1174-1180. Retrieved from http://www.lrec-conf.org/proceedings/lrec2010/pdf/106_Paper.pdf

Corder, S. P. (1981). *Error Analysis and Interlanguage*. London: Oxford University Press.

Craciunescu, O., Gerding-Salas, C., & Stringer- O'Keefe, S. (2004). Machine Translation and Computer-assisted Translation: a New Way of Translating?. *Translation Journal*, 8(3). Retrieved from http://www.translationjournal.net/journal

Creswell, J. W. (2012). *Planning, Conducting, and Evaluating Quantitative and Qualitative Research (4th ed.).* Boston: Pearson Education.

Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided Analysis. *System*, 26, 163-174. Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/S0346251X98000013

Delavenay, E. (1960). *An Introduction to Machine Translation*. Great Britain: The Camelot Press Ltd. Southampton.

Edwin, S.K.S. (2013). *Practical English*. Malaysia: Edward Yoshioka.

Eftekhar, M. & Nouraey, P. (2013). Commercial Translation Error Analysis: A Case Study of Iranian Products. *Linguistics and Literature Studies*, 1(2), 55-60. Retrieved from http://www.hrpub.org/download/201309/lls.2013.010201.pdf

Elliot, D., Hartley, A., & Atwell, E. (2004). A Fluency Error Categorization Scheme to Guide Automated Machine Translation Evaluation. *Computer Science*, 32(65), 64-73. Retrieved from http://www.comp.leeds.ac.uk/eric/elliott04amta.pdf

Esteki, A. (2010). *What is to be considered as a Unit of Translation?.* Retrieved from http://www.translationdirectory.com/articles/article2124.php

Gass, S.M. & Selinker, L. (1994). *Second Language Acquisition: An Introductory Course (3rd ed.).* New York: Routledge.

Gaurdat, H., Hughes, R. & Michael, J. (1996). *Towards Better English Grammar*. Selangor: Penerbit Fajar Bakti Sdn. Bhd.

Google Translate. (n.d). *Inside Google Translate*. Retrieved from http://translate.google.com.my/about/intl/en_ALL/

Graham, Y. (2015). Improving Evaluation of Machine Translation Quality Estimation. *Computational Linguistics*, 7, 1804-1813. Retrieved from http://www.aclweb.org/anthology/P15-1174.pdf

Hadis, G. & Mahmood, H. (2016). A Comparative Study of Google Translate Translations: An Error Analysis of English-to-Persian and Persian-to-English Translations. *English Language Teaching*, 9(3), 13-17. Retrieved from http://dx.doi.org

Handschuh, C. (2013*). Human Translation vs. Machine Translation: A Triumph of Nuance over Nonsense?*. Retrieved from http://dculs.dcu.ie/dcu-language-services-news/human-translation-vs-machine-translation

Hatim, B. & Munday, J. (2004). *Translation: An Advanced Resource Book*. London: Routledge.

Henry, A. (2014). *Most Popular Language Translation Tool: Google Translate*. Retrieved from http://lifehacker.com/five-best-language-translation-tools-1634228212/1636028816/+alanhenry

Hidayatul Akmal. (2007). *Kanser usus tidak kenal keturunan*. The New Straits Time Press (M) Berhad. Retrieved from http://www2.hmetro.com.my/myMetro/articles/Kanserususti dakkenalketurunan/Article/article_print

Hossein, V. & Saleh, D. (2012). A Study of Translation Problems of Tourism Industry Guidebooks: An Error Analysis Perspective. *Foreign Language Teaching and Research*, 7(1), 71-82. Retrieved from http://jfl.iaun.ac.ir/article_4206_1.html

House, J. (1981). *A Model for Translation Quality Assessment*. Hamburg: Gunter Narr Verlag.

House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Translators' Journal*, 46(2), 243-257. Retrieved from www.erudit.org

House, J. (2015). *Translation Quality Assessment Past and Present*. New York: Routledge.

Hutchins, W. J. (1995). Machine Translation: A Brief History. In Koerner, E. F. K., & Asher, R. E (Eds.), *Concise history of the language sciences: From the Sumerians to the Cognitivists*, 431-445. Oxford: Permagon Press. Retrieved from http://www.aymara.org/biblio/mtranslation.pdf

Hutchins, W. J. & Somers, H. L. (1992). *An Introduction to Machine Translation*. London: Academic Press. Retrieved from http://www.hutchinsweb.me.uk/IntroMT-TOC.htm

Jakobson, R. (1959). On Linguistic Aspect of Translation in Venuti, L. *The Translation Studies Reader* (Ed.) (2000), 113- 120. London: Routledge

Jia, Y., Muyun, Y., Junmao, M., Dongqing, X., Tiejun, Zhao., & Sheng, L. (2012). How Good is Web Based MT: A Case Study on Paper Title Translation. *International Conference on Asian Language Processing*, 40, 213-216. Retrieved from http://www.computer.org/

Juan, Du. (1994). *Machine Translation for a Group Decision Support System*. Master of Science. Department of Information Systems and Computer Science. National University of Singapore

Kamus Dewan Edisi Keempat. (2013). Kuala Lumpur: Dewan Bahasa dan Pustaka.

Karami, O. (2014). *The Brief View on Google Translate Machine*. Retrieved from http://ceur-ws.org/Vol-867/Paper17.pdf

Keshavarz, M.H. (1999). *Contrastive Analysis and Error Analysis*. Tehran: Rahnama Press. Retrieved from http://bookyar.net/?paperno=142803

Komeili, Z., Jalil, F.H., & Ali, R. (2011). An Investigation of the Translation Problems Incurred by English-to-Persian Machine Translation:"Padideh, Pars, and Google Softwares".*Procedie-Social and Behavioral Sciences*, 28, 1079-1082. Elsevier. Retrieved from www.scinecedirect.com

Koponen, M. (2010). Assessing Machine Translation Quality with Error Analysis. *Proceedings on translation and Interpreting Studies*, 4, 1-12. Retrieved from http://www.sktl.fi/@Bin/40701/Koponen_MikaEL2010.pdf

Koponen, M. & Salmi, L. (2015). On the Correctness of Machine Translation: A Machine Translation Post-editing Task. *The Journal of Specialised Translation*, 23, 118-136. Retrieved from http://www.jostrans.org/issue23/art_koponen.pdf

Kulwindr Kaur Sidhu. (2006). *Translating Scientific Texts: English to Malay*. Kuala Lumpur: University of Malaya Press.

Larson, M. (1998). *Meaning-Based Translation: A Guide to Cross-Language Equivalence (2nd ed.)*. New York : University Press of America.

Lauer, B. (1995). Abstracts in German Medical Journal: A Linguistic Analysis. *Information Processing & Management*, 31(5), 769-776. Retrieved from http://eric.ed.gov/?id=EJ513775

Lindner, M. (2005). *English Language & Composition*. Selangor: Printmate Sdn. Bhd.

Matsuzaki, T., Fujita, A., Todo, N., & Arai, H. (2016). *Translation Errors and Incomprehensibility: A Case Study Using Machine-Translated Second Language Proficiency Tests*. Retrieved from http://www.lrecconf.org/proceedings/lrec2016/pdf/921_Paper.pdf

Melby, A.K. (1987). On Human-machine Interaction in Translation. In S.Nirenburg (Ed.), *Machine Translation*: *Theoretical and Methodological Issues*, 145-154. United Kingdom: Cambridge University Press.

Mogahed, M. (2011). To Use or not to Use Translation in Language Teaching. *Translation Journal*, 15(4). Retrieved from http://translationjournal.net/journal/58education.htm

Mohamad Nor Amin & Naimah Abdullah. (2011). Analisis Sintaksis Melalui Terjemahan Google Melayu-Arab in Noor Ida, Norizah & Aini Aziz (Ed). *Globalisasi Melalui Terjemahan: Pemangkin Kecermelangan Ilmu dan Teknologi*. Kuala Lumpur: Persatuan Penterjemah Malaysia.

Munday, J. (2008). *Introducing Translation Studies: Theories and Applications*. London: Routledge.

Newmark, P. (1988). *A Textbook of Translation*. London: Prentice Hall.

Newmark, P. (1991). *About Translation.* Great Britain: Multilingual Matters Ltd.

Nida, E. A. & Taber, C. R. (1964). *Towards a Science of Translating: With Special Reference to Principles and Procedures involved in Bible Translating*. Leiden: E. J. Brill.

Nida, E. A. & Taber, C. R. (1969). *The Theory and Practice of Translation*. Leiden: E. J. Brill.

Norhazlina Husin. (2009). Penyesuaian dalam Penterjemahan Brosur Bank Malaysia: Satu Kajian Kes. *Jurnal Penterjemah*, 24(2), 121-142.

Norizah Ardi, Midiyana Mohamad, Rozaimah Rashidin & Zuraidah Jantan. (2009). *Pengantar Prinsip Penterjemahan*. Selangor: August Publishing Sdn. Bhd.

Ping, K. (2009). Machine Translation. In Baker, M. & Saldanha, G. (Ed.). *Encyclopedia of Translation Studies. (2nd ed.).* New York: Routledge.

Popovic, M. & Burchardt, A. (2011). *From Human to Automatic Error Classification for Machine Translation Output*. Retrieved from www.dfki.de

Popovic, M.& Ney, H. (2011). Towards Automatic Error Analysis of Machine Translation Output. *Computational Linguistics*, 37(4), 657-688. Retrieved from http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_a_00072

Presada, D. & Badea, M. (2014). The Effectiveness of Error Analysis in Translation Classes. A Pilot Study. *Porta Linguarum*, 22, 49-59. Retrieved from http://www.ugr.es/~portalin/articulos/PL_numero22/4%20%20DIANA%20PRESADA.pdf

Reiss, K. (1971). Type, Kind and Individuality of Text. (Kitron, S., Trans.) in Venuti, L. *The Translation Studied Reader (Ed.)* (2000), pp. 160-170. London: Routledge.

Rensburg, A., Snyman, C., & Lotz, S. (2012). Applying Google Translate in a higher education environment: Translation products assessed. *Southern African Linguistics and Applied Language Studies*, 30(7), 511-524. Retrieved from http://www.tandfonline.com

Sager, J. (1983). *Quality and Standards- The Evaluation of Translations*. *The Translator's Handbook*. London: ASLIB.

Sager, J. (1994). *Language Engineering and Translation: Consequences of Automation*. Amsterdam: John Benjamins Publishing Company.

Shuttleworth, M. & Cowie, M. (1997). *Dictionary of Translation Studies*. Manchester, UK: St. Jerome.

Stymne, S. & Ahrernberg, L. (2012). On the Practice of Error Analysis for Machine Translation Evaluation. *International Conference on Language Resources and Evaluation*, 1785-1790. Retrieved from https://www.semanticscholar.org/paper

Talal Mohammad. (2012). *An Evaluation of the Use of Machine Translation among Arab Students.* Master's Dissertation. Faculty of Languages and Linguistics. Kuala Lumpur: University of Malaya.

Temnikova, I. (2010). Cognitive Evaluation Approach for a Controlled Language Post-editing Experiment. *LREC*, 3485-3490. Retrieved at http://www.lrec-conf.org/proceedings/lrec2010/pdf/437_Paper.pdf

Tengku Sepora Tengku Mahadi & Mohamad Abdou. (2006). *Text-wise in Translation*. Malaysia: Prentice Hall.

Toma, P. (1959). An Operational Machine Translation System in Brislin, R, W. *Translation Applications and Research (Ed.)* (1976), 248-257. United States: Gardner Press Inc.

Trask, R.L. (1993). *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge.

Venuti, L. (2000). *The Translation Studies Reader*. New York: Routledge.

Vilar, D., D'Haro, L.F., & Ney, H. (2006). Error Analysis of Statistical Machine Translation Output. *International Conference on Language Resources and Evaluation*, 697-702 Retrieved from http://www-i6.informatik.rwth-aachen.de

Vinay, J. P., & Darbelnet, J. (1958). A Methodology of Translation. (Sager, J. C., & Hamel, M. J., Trans.) in Venuti, L. *The Translation Studies Reader (Ed.)* (2000), 84-93. London: Routledge.

Wilks, Y. A. (1972). *Grammar,Meaning and The Machine Analysis of Language*. London: Routledge.

Williams, M. (2004). *Translation Quality Assessment: An Argumentation Centred Approach*. Ottawa: University of Ottawa Press. Retrieved from https://www.erudit.org/revue/ttr/2004/v17/n2/013278ar.pdf