

Education Quarterly Reviews

Sutrisno, Adi. (2020), The Accuracy and Shortcomings of Google Translate Translating English Sentences to Indonesian. In: Education Quarterly Reviews, Vol.3, No.4, 555-568.

ISSN 2621-5799

DOI: 10.31014/aior.1993.03.04.161

The online version of this article can be found at:
<https://www.asianinstituteofresearch.org/>

Published by:
The Asian Institute of Research

The *Education Quarterly Reviews* is an Open Access publication. It may be read, copied, and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

The Asian Institute of Research *Education Quarterly Reviews* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of education, linguistics, literature, educational theory, research, and methodologies, curriculum, elementary and secondary education, higher education, foreign language education, teaching and learning, teacher education, education of special groups, and other fields of study related to education. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Education Quarterly Reviews* aims to facilitate scholarly work on recent theoretical and practical aspects of education.



ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

The Accuracy and Shortcomings of Google Translate Translating English Sentences to Indonesian

Adi Sutrisno¹

¹ Faculty of Cultural Sciences, Universitas Gadjah Mada. Email: adisutrisno@ugm.ac.id

Abstract

Google Translate is a free and practical online translation service that allows millions of people around the globe to translate words, phrases, sentences, and paragraphs into an intended target language. However, in 2015, some Google Translate users in Indonesia, filed complaints, asserting that the machine was often inaccurate, speculating that it could only translate languages at the micro-level of words and phrases, rather than complete sentences or paragraphs. This research works to examine the accuracy as well as the shortcomings of Google Translate, in the context of English to Indonesian translations, in order to critically engage the complaints made by Google users. For the purpose of this study, 80 English sentences were translated using Google Translate and assessed for accuracy using a table adapted from Memsorce criteria. Both the original sentences and their translated versions were analyzed using a sentence pair matrix to determine the machine's failings and areas for improvement. The results challenged those initial speculations which suggested Google Translate is only effective with words and phrases. On the contrary, Memsorce proved to be a useful tool in demonstrating a reasonable level of accuracy, accurately translating 60.37% of Indonesian-English sentences and vice versa.

Keywords: Google Translate, Memsorce, Statistical Machine Translation, English-Indonesian Translation, Sentence-Pair Matrix

Introduction

There has been an upsurge of interests in machine translators within the last two decades, which has motivated the providers behind machine translations to improve their services. Thus far, the quality of many of these automated translations has left many users unsatisfied. *Google Translate*, for instance, remains a work-in-progress (Barreiro, et al 2014) despite its popularity as the most widely used machine translator in the world (Aitken and Balan, 2009; Butler, 2011).

As demonstrated by Aitken and Balan (2009) and Patil and Davies, (2014), *Google Translate* works well with some languages, but is less effective with others. As Patil and Davies (2014) reported in their research article on the translation of medical terms using *Google Translate*, the accuracy of translation varied from language to language. In their research, Patil and Davies (2014) tested ten commonly used medical statements. The statements were translated from English into 26 different languages using *Google Translate*. The translated statements were then sent to native speakers from each respective language, who were asked to translate the statement back into English and send it back to the researchers. The returned English phrases were compared with the originals and

assessed for meaning. The translations which did not make sense or were factually incorrect were marked as wrong, while minor grammatical errors were allowed. The researchers concluded that when the statements were translated into Eastern European languages, the accuracy of the translation reached 62%; and when the statements were translated into Western European languages, the accuracy was higher, around 74%. However, when the phrases were translated into African languages, the translations only reached a 45% accuracy level. Similarly, when translated into Asian languages, they were 46% accurate.

Patil and Davies (2014), reported some serious errors in the translations such as (a) “Your child is fitting” translated in Swahili to “Your child is dead.” (b) “Your husband has the opportunity to donate his organs” translated into Polish to “Your husband can donate his tools.” (c) “Your husband had a cardiac arrest” translated to Marathi as “Your husband had an imprisonment of heart.”; and “Your wife needs to be ventilated” translated to Bengali as “Your wife wind movement needed.” In their conclusion, Patil and Davies advised against using *Google Translate* to translate medical terms.

In relation to the findings of Patil and Davies (2014), the present writer intends to examine *Google Translate*'s accuracy in translating English sentences to Indonesian and to analyze the potential advantages and drawbacks of such translations. In light of these issues concerning Indonesian-English translations, I feel compelled as a translation studies researcher, to explore the extent that *Google Translate* was criticized to be ‘only good at transferring words from one language to another, but not sentences’ (see ONLINE translator and ENGLISHINDO.COM) and its subsequent framing as unsuitable for use, especially regarding text which has a significant cultural content.

On a level of pedagogical one, my curiosity grew as my students studying Indonesian language and culture at Inculs shared their personal experiences with *Google Translate* and its shortcomings, particularly when the machine was used to translate a number of phrases with significant grammatical complexity. Inculs or Indonesian Culture and Language Studies is a language institution specializing in the study of Indonesian language and culture. It belongs to Universitas Gadjah Mada of Indonesia. The students are from various countries from all over the world.

Before embarking on further research on how *Google Translate* operates in Indonesian contexts, I had to first consider what *Google Translate* is. It is a machine which “gathers a huge database of human-revised translations of millions of documents” (Grajales, 2015). This is in fact one of the strengths of *Google Translate* as a ‘statistical-based machine’, especially when it is used to translate languages having similar grammatical system. As Aitken and Balan (2009); Grajales, (2015); Patil and Davies (2014) reported, it can reach an accuracy level of more than 70% in English – European language pairings.

According to Osborne, et al (2006:17-24), a high level of accuracy between languages with similar grammatical systems is enabled by the operational system of the machine which is believed to be using parallel corpora. Theoretically, a statistical machine translation using arallel corpora such as translating English sentences to Portuguese, Indonesian to Malay, or Dutch to German, and vice versa will be easily executed by *Google Translate*. In contrast, translating languages made up of different language systems will not be as accurate

This research is intended to carry out a preliminary exploration of issues concerning the level of accuracy of *Google Translate* in translating sentences from English to Indonesian; with a goal of shedding light on the potential advantages and drawbacks of *Google Translate* when used to translate English sentences to Indonesian.

How *Google Translate* Works

Google Translate is a machine translation tool which uses statistics as the basis of its operation (Aitken and Balan, 2009; Butler, 2011; Barreiro, et al., 2012; Graesser et al, 2014). As a statistical machine translator, *Google Translate* uses linguistic modeling, statistical decision theory, and matching probabilities. (Ney, 1995:107-119; Koehn et al, 2003:48-54) as the basis for its translation processing. With this advantage, *Google Translate* has the capability, with the proper data sources, to translate sentences from any source language to any target language.

Since *Google Translate* uses statistics as the basis of its operation, the higher the degree of similarity between the initial source language and the target language, the higher the degree of accuracy the translator might achieve (see Koehn et al, 2003:48-54).

To understand the work of a statistical machine translation, one needs to be familiar with Bayes Theorem, proposed by Brown et al. (cited in Okpor, 2015), theorizing that any sentence in one language can potentially be the translation of any sentence in another and “the most appropriate is the translation that is assigned the highest probability by the system”. It means that no matter how different a target language is from a source language, a translation between the two can still be generated; while the chance of being accurate is affected by the probability distribution function which is commonly indicated by $p(e|f)$ and the language model, by $p(e)$. The probability distribution function ensures that the machine translation system produces a target hypothesis corresponding to the source sentence; while the language model ensures a grammatically correct output. (Okpor, 2015). This is how *Google Translate* facilitates translations.

In relation to the workings of the machine, Osborne et al (2006:17-24) contends that a statistical machine translation uses bilingual text corpora which operates systematically to translate a source language to a target language. An example is given to illustrate how a statistical machine translation like *Google Translate* works. In the following example a French sentence is translated into English.

1. *Ces gens ont grandi, vecu et ouvree- des dizaines d'annees dans le domain agricole.*
2. Those people have grown up, lived and worked many years in a farming district.

The English sentence is translated to French using the bilingual text corpora as follows:

<i>Ces gens ont</i>	Those people have
<i>gens ont grandi</i>	people have grown up
<i>ont grandi</i>	have grown up
<i>grandi, vecu</i>	grown up, lived
<i>, vecu et</i>	, lived and
<i>vecu et oeuvre</i>	lived and worked
<i>et oeuvre des dizaines d'oeuvre</i>	and worked many
<i>oeuvre des dizaines d'annees dizaines</i>	worked many years
<i>des dizaines d'annees dans</i>	many years in
<i>anes dans le</i>	years in a
<i>le domain agricole</i>	a farming district
<i>domaine agricole.</i>	farming district

In order to produce a translation, first the machine parses the source language into segments, identifying the match in the target language. After finding the best distribution, the machine determines the best match for the sentence.

As one may observe, the English translation of the sentence ‘fits’ the French version in many respects. While French and English differ in some grammatical respects, such as noun-adjective word order, and are gender marked differently, as Western European languages, they share many similar characteristics (Patil and Davis, 2014). For example, the similar word order of the two languages used in this translation provides ease in the process of pairing, even between segments within the sentences (Osborne et al, 2006).

As table 1 below illustrates, sentence pair alignment occurs systematically at the word-level in the previous example of French-English translation, although there is a change on noun-adjective order at the end of the sentence.

Table 1: The sentence pair is: Those people have grown up, lived and worked many years in a farming district.

	Those	people	have	grown	up	,	lived	and	worked	many	years	in	a	farmin	o	district	.
<i>Ces</i>	█																
<i>gens</i>		█															
<i>ont</i>			█														
<i>grandi</i>				█	█												
<i>,</i>						█											
<i>vecu</i>							█										
<i>et</i>								█									
<i>ouvree</i>									█								
<i>des</i>										█							
<i>dizaines</i>											█						
<i>d'</i>												█					
<i>annees</i>													█				
<i>dans</i>														█			
<i>le</i>															█		
<i>domain</i>																█	
<i>agricole</i>																	█
<i>.</i>																	

Here we see that the translation of languages of the same language family — in this case from French to English — forms an oblique pattern when viewed in the pairing table, moving from the top left to the bottom right. This diagonal arrangement reflects the workings of a one-to-one relation system in the translation of the text from the source language into the target language. In the context of languages which do not have a similar grammatical structure, the translation of the language pair will likely show a different, non-specific pattern of flow if plotted in the matrix table.

The Popularity and Usefulness of Google Translate to Indonesian Audiences

Despite the potential pitfalls, as briefly outlined above, *Google Translate* remains one of the most popular translation machines in Indonesia. As reported by the Engineering Director of *Google Translate* MacDuff Hughes, Indonesian is one of the top ten languages in the world used in *Google Translate* after English, Spanish, Arabic, Russian, and Portuguese (Liputan 6, 2016).

According to Julie Cattiau, Product Manager of *Google Translate*, most Indonesians use *Google Translate* in order to understand and share information on the internet (Liputan 6, 2016). Other uses include understanding unfamiliar languages when traveling abroad and communicating with business counterparts. As explained by Napitupulu (2017:15-23) *Google Translate* is even used by students at the Methodist University of Indonesia, Medan, North Sumatra (Napitupulu, 2017:15-23) to translate their undergraduate abstracts from Indonesian to English. Similarly, Farlin (2015) claimed that Indonesian students with a limited grasp of English often translate their graduate paper abstracts into English using *Google Translate*. Thus it can be surmised that the popularity and usefulness of *Google Translate* in Indonesia is a real and significant phenomenon, for that reason, it is worthwhile to investigate the nuances of *Google translate* in Indonesia and examine its sociocultural impact.

Indonesian Language at a Glimpse

Indonesian, which belongs to the Austronesian language family (Sneddon, 2003) is the official language of Indonesia, and is now spoken by more than 200 million people throughout the country. The language is actively spoken by majority of people, and thus functions as a unifying language for all the people living throughout

Indonesia, scattering 34 provinces, from Aceh to Papua. There are more than 700 indigenous local languages which are used at regional and district levels. Thus, the majority of Indonesian people are polyglots.

Indonesian is composed of words from various local languages, especially Javanese; and today is even influenced by English words and dialects due to globalization and encroaching elements taken from the mass media, internet sources, or interaction with people from other countries. Thus, there are many new words that appear in Indonesian which have been borrowed or adapted from other languages. For example: *action*, *riset* (research), *komunikasi* (communication), *promosi* (promotion), *provokasi* (provocation), *probabilitas* (probability), *kontinuitas* (continuity), *regularitas* (regularity), and so on.

In terms of its structure, Indonesian is similar to English, as it begins with a Subject followed by a Verb and an Object or SVO. This similarity in word order within a sentence makes it easy to translate from English into Indonesian, and vice versa. However, there are some cases which might cause problems during the translation process, namely: (1) Indonesian does not have verb tenses; (2) Indonesian is rich with prefixes, confixes, and infixes, including the prefix ‘di-’ and ‘me’, which indicate passive and active voices; (3) Unlike English, Indonesian does not have the gerund or participle tense; (4) Lastly, Indonesian does not have subjunctive verbs.

Given that Indonesian does not have verb tenses, it is unsurprising to hear that some Indonesian English-language learners struggle to produce the correct tenses in English sentences. A sentence like “Borobudur is built in the ninth century by Cailendra dynasty” is just one of the many errors which might be produced by Indonesian students in their initial stages of learning English. This problem is related to the difficulty of using verb tense which is not in existence in Indonesian language. Likewise, those who study Indonesian might find it difficult to anticipate this tense difference and navigate sentences accordingly. They often memorize temporal adverbs in Indonesian, knowing that when they translate English sentences with certain tenses, they can pair the verbs with temporal adverbs that match the English verb tenses. It is worthwhile to explore how *Google Translate* addresses this discrepancy, and how accurate its results are, especially when tasked with translating English sentences with complex verb tenses to Indonesian.

It is also noteworthy how *Google Translate* anticipates the difficulty of translating passive voice from English to Indonesian, as the passive voice in English and Indonesian differ both morphologically and semantically. In Indonesian, passive voice can be identified by the presence of prefix *di* + root verb or *ter* + root verb. For instance :

- (1) Daging itu **di**iris dengan hati-hati
- (2) **Di**a **ter**iris jarinya saat memotong-motong bawang merah.

The difference between prefix **di-** and **ter-** is that the prefix **di-** indicates there is intention in the action; or you do it on purpose; while prefix **ter-** indicates there is no intention in the action; or you do not do it on purpose. You cannot use prefix **me-** because **me-** is only used in active voice. Therefore you cannot say

- (3) Daging itu mengiris dengan hati-hati.
The meat cuts with care
- (4) Dia mengiris jarinya saat memotong-motong bawang merah.
He cuts his finger when chopping onions

Sentence (3) does not make sense. Meat can not do the cutting; while sentence (4) means He cut his finger when chopping onions (and he cut it with intention, on purpose).

Another aspect which may cause problems in translations is that Indonesian does not have ‘gerundial form’ or ‘present participles’ which take the same form in English, namely, the ‘-ing form’. Thus, the phrases (1) sleeping dog and (2) sleeping pill have exactly the same form (-ing) and position (before noun) but different functions. (1) A sleeping dog is a “dog which is sleeping”; but (2) a sleeping pill is not a “pill which is sleeping”, as this does

not make sense. The correct meaning is a “pill for sleeping”, a gerundial form. However, in Indonesian there is no such form.

Additionally, Indonesian does not have the same subjunctive form as in English. In English the subjunctive form is indicated by the presence of constructions such as “If I were you, I would marry her” or “Had I had that opportunity, I would have taken it”. Because there are no verb tenses in Indonesian, it becomes difficult to understand subjunctive tense for the Indonesian English-language learners.

The different grammatical structures of Indonesian and English challenge whether *Google Translate* can accurately and effectively translate Indonesian sentences into English and vice versa, or if is just a rudimentary tool appropriate solely for extracurricular and recreational purposes. This research is meant to determine how accurate the machine is; what advantages it poses, and what potential shortcomings or problems manifest which need to be addressed.

Research Method

Data samples

As cited in the aforementioned discussion, my research was inspired by witnessing Universitas Gadjah Mada’s international students’ difficulties in translating English to Indonesian and vice versa, using *Google Translate*. experienced by international students who take English-Indonesian translation course at Inculs at the faculty of cultural sciences Universitas Gadjah Mada of Indonesia. In the light of the grammatic and semantic construction disparities, difficulties of English-Indonesian translation frequently occurred on syntactical and semantical level, particularly in contexts, of (1) passive voice, (2) conditional sentences, (3) sentences containing verb tenses (4) and sentences containing elements of gerund and participles. As the first study to consider *Google Translate* on a granular level, this study does not purport to ‘explain’ all challenges of English-Indonesian translation. As an exploratory study, the scope of the data used in this paper is limited to English sentences which contained the four grammatical elements above. The total number of sentences examined was 4 x 20 sentences, or 80 sentences. Though these four elements are the primary focus, word selection, which offers additional nuance and challenges, is also discussed in this paper. This is because an author’s intentional word selection or diction often has a context-specific meaning which is sometimes not recognized by the readers.

Research Procedure

As an exploratory study into the difficulties of English-Indonesian translation, the methodology used combined a comparative and descriptive approach, combining *Google Translation* and ‘human’ translation. In the first instance, I selected 4 grammatical cases which often cause problems to Indonesian English-language learners. They are ‘verb tenses’, ‘gerund and participle’, ‘passive voice’, and ‘conditional sentences’. There are 20 sentences with each case, so there are 80 sentences altogether.. Each sentence was translated using *Google Translate*. These four grammatical cases were chosen as the focus of this examination because (a) tenses do not exist in the Indonesian language, and thus might create problems in the translation process; (b) gerunds and participles have the same form in English, signified by forming a Verb+ing, which might cause confusion for Indonesian English-language learners; (c) passive voice is English and Indonesian have different constructions to a certain extent, and it might make it difficult for those learning Indonesian; and (d) conditional sentences are expressed using certain verb tenses which are not in existence in Indonesian and thus might create difficulty in the translation process.

I then requested 4 translation experts to translate all of the 80 sentences (human translation) as one way of exploring the similarity between the *Google Translate* results and human translation.

The next stage of the research was to ‘score’ the results of *Google Translation* using ‘translation closeness metric on the basis that “the closer a machine translation is to a professional human translation, the better it is” (Pepineni et al; 2012).

The following is a 'translation closeness metric' adapted from Memsorce criteria to determine the translations' accuracy level.

No	Level of Accuracy	Description of accuracy
1	Match 100	The translation results are identical to human translation by expert translations, with no editing required.
2	Match 85 – 95	The translation is near perfect, but requires a bit of editing by translation experts.
3	Match 50 – 75	The translation results require improvements to the words used, and must be edited by translation experts
4	Match 00 – 49	The translation results are still incomplete and tend not to match the source language.

(blog.memsorce.com)

After the level of accuracy was identified, the shortcomings of Google Translate were analyzed using Osborne's (2010) sentence-pair matrix.

Research Results and Discussion

The Level of Accuracy of Google Translate translating English Sentences to Indonesian.

Based on evaluations using the 'translation closeness metric' which is commonly applied by Memsorce (2012), when translating English sentences into Indonesian *Google Translate* reached an accuracy level of more than 60%. It is significantly worse than human translation, but much better than when African and other Asian languages were translated, as reported by Patil and Davies (2014), which reached 45% and 46% consecutively. The individual achievement in the degree of closeness for each item in this research was as follows: (1) The translation of sentences having certain tenses reached 62.5% accuracy; (2) The translation of conditional sentences reached 58.5%, (3) sentences containing Gerund and Participle 63.5%, (3) and passive sentences reached 57 % accuracy.

One example of an English sentence written in the future tense, translated to Indonesian is shown as follows:

01. They will plant roses in the garden.
02. Mereka akan menanam bunga mawar di taman.

Table 2: The sentence pair *They will plant roses in the garden.*

	Mereka	akan	menanam	bunga	mawar	di	taman
They							
will							
plant							
roses							
in							
the							
garden							

Table 2 shows the accuracy of the translation reflected in the word pairings which correspond to each other in a manner similar to the aforementioned French-English translation. The sentence pair in table 2 also reveals a diagonal line from top left to bottom right, meaning that the pattern of the word order of both languages is similar.

Many, though not all, of the sentences with other verb tense forms in their grammatical constructions were also translated correctly. Of the 20 sentences containing past tense, past perfect tense, past continuous tense, present perfect tense, future tense, 37.5% were translated inappropriately, while 62.5% were translated correctly.

The Shortcomings of Google Translate in Translating English Sentences to Indonesian.

The results of this study contradict views that Google Translate is limited to translating words or phrases in the context of English-Indonesian translation. In fact, the results of this study indicate that the *Google Translate* has a capability of translating sentences although none of the ‘machine-translated’ sentences could match the human translator in terms of the quality of translation.

The issues facing *Google Translate* in the context of English-Indonesian translation is related, for the most part, to sentences having a different point of view between the source language and the target language. By this I mean there are some cases where translators intentionally generate a change in the point of view of the message without altering its meaning and without generating an unnatural feeling in the reader of the target text. In this respect, *Google Translate* is unable to do that. The following is the example of the shortcomings of Google Translated translating sentences with different points of view.

Sample 2.

Sentence to translate	Google Translate	Human Translation
He dropped his wallet.	<i>Dia menjatuhkan dompetnya.</i>	<i>Dompetnya terjatuh</i>
	He dropped wallet+his	Wallet+his dropped

Table 3: The comparison between Google Translate and Human Translation

Table 3A: *Google Translate*

	Dia	Menjatuhkan	Domet	nya	.
He					
dropped					
his					
wallet					
.					

Table 3B: Human Translation

	Dompot	nya	terjatuh		.
He					
dropped					
his					
wallet					
.					

The data in table 3A shows a sentence pair which has been translated through *Google Translate*. In the pairing image we see the word order making a diagonal line, from top left to bottom right, with little shift of box formation on column three and column four. This diagonal line illustrates a similar trend formed between the target language and the source language; while the shifting formation of the box in columns three and four that do not make the diagonal line indicates there is different construction in that section.

Data in table 3B shows the sentence pair translated through human processes. The Configuration of dark boxes which shows the trend of the sentence moves from bottom-left to top-right. The pattern is different from data in table 3A. And in terms of acceptability, the translation as seen in 3B is more appropriate compared to the translation in 3A. It is congruent with what Papineni (2002) stated, that the quality of machine translation will never surpass that of human translation.

One of the advantages of human translators compared to statistical machine translators is that human translators have a sense of the language, or modulation while a statistical machine translator do not yet have this capability. When a string of sentence of a source language is unable to be verbatim translated to a target language, a change of perspective, or structural reconfiguration, is called for to enable a target language to convey the proper meaning. This situation is commonly termed as modulation. According to Hardin and Picot (1990:21) modulation is a change in perspective that allows a translator to express the same phrase or sentence in different ways. Notice that the sentence “He dropped his wallet” was translated to *Dia menjatuhkan dompetnya*. In Indonesian this sentence means “He intentionally dropped his wallet”. If one does not intend to drop the wallet, the order is reversed: *Dompetnya terjatuh*, which literally means “His wallet was dropped”. This nuance proves that the point of view or perspective of the source language cannot be transferred into the target language by machine translation alone and must be checked by the human translator.

Other similar sentences which were not translated appropriately by *Google Translate*:

05. The pizza was so hot that he burnt his tongue.
06. She cut her finger when chopping onions.
07. I got something in my eye.
08. I hit my head on the wall

The above sentences were not translated appropriately by *Google Translate* because of the different way of expressing active and passive voices in the source and target languages. When translating the source language, *Google Translate* followed the pattern of the source language so that the translation still has the sense of the source language. As a consequence, the translation was inappropriate.

Take a look at the translation using *Google Translate* below:

05. *Pizza itu begitu panas bahwa dia membakar lidahnya.*
06. *Dia memotong jari-jarinya ketika memotong bawang.*

The translation results showed that there is something missing in the proposition. It is a matter of how the proposition is expressed in the target language. The clause (05) “... he burnt his tongue” when it is translated as it is, is interpreted as an act of ‘deliberately burning his tongue’, while the intent of the sentence is actually “his tongue burns.” The same is true in clause (06) “He cuts his fingers” which is interpreted as an act of “deliberately cutting his fingers”. These examples prove that the machine translation is currently unable to transform the tacit meaning contained in the source language to the target language.

Similar cases can be seen in sentence (07) and (08). Sentences that read (07) “I have something in my eyes” and (08) “I hit my head on the wall” would generate a change in a point of view of the message. In Indonesian, the subject in sentence (07) is ‘my eyes’ and in sentence (08) the subject is “my head”. In Indonesian it would be appropriate and correct if those phrases become the focus of attention in the sentence construction, therefore they are placed as subject of the sentence or antecedent. The machine is unable to facilitate these subtle grammatical shifts. Therefore, a human translator must create a change in perspective so the translated sentences convey the proper meaning. Hence, instead of translating to (07) “*Saya mempunyai sesuatu di mata saya*” and (8) “*Saya memukulkan kepala saya di tembok*”, the human translator translated them as (07b) *Mata saya kemasukan sesuatu* dan (08b) *Kepala saya terbentur tembok*.

Word Selection in the Target Language

Another problem experienced by *Google Translate* users when translating English sentences to Indonesia is the accuracy of context-based word selection. The inaccuracy in word selection can result in a shift of meaning, both at the micro level (invert), namely in the level of words or phrase and macro (convert) or in the level of wider context, namely meaning as a whole. The meaning of sentences that contain local errors can still be generated to

a certain degree; while the meaning of sentences having global errors will result in significant deviations of meaning.

The quality of *Google Translate* as a statistical machine translation is affected by the accuracy of the inputted diction. If the words used in translation are contextually correct, the result is quality translations; however, if the words used in this translation are contextually incorrect, the translation will not be of sufficient quality.

The results of this study indicate that *Google Translate* in general has been able to use words appropriately given their context. The only shortcoming the *Google Translate* has, in terms of word selection, is mainly on words having multiple functions such as “that” and “used”.

Examples

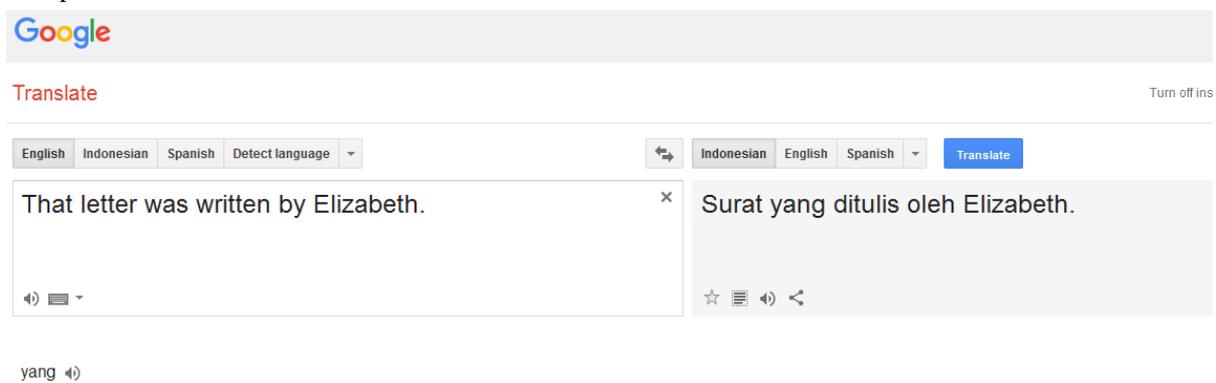


Figure 1: The translation of the sentence “That letter was written by Elizabeth”

The translation of the source language distorts the grammatical unit. The input, which in the source language was a sentence, is transformed into a clause. It is mainly because of the machine’s inability to identify the function of the word ‘that’ in the sentence. The sentence “That letter was written by Elizabeth” when translated by the machine becomes “The letter that was written by Elizabeth” in the target language. Thus, it changes the unit, from a sentence to a clause.

The word ‘that’, and other demonstratives, have the potential to cause errors in machine translation. ‘That’ can have multiple functions. It can function as (1) ‘demonstrative adjective’ as in the sentence “That test was hard”, (2) ‘demonstrative pronoun’ as in sentence “That was hard”, (3) ‘adverb’, as in “The test was not that bad”. The word ‘that’ can also introduce ‘restrictive relative clause’ as in “The test that she took was hard.”

Another example that is often mistranslated in *Google Translate* is the phrase ‘used to’. In general *Google Translate* identifies the phrase used to as the word "used" so that the translation results experience a global error, or there is a deviation of meaning that is too far from the original source for comprehension. Notice the translation of the phrase “used to” in the following:

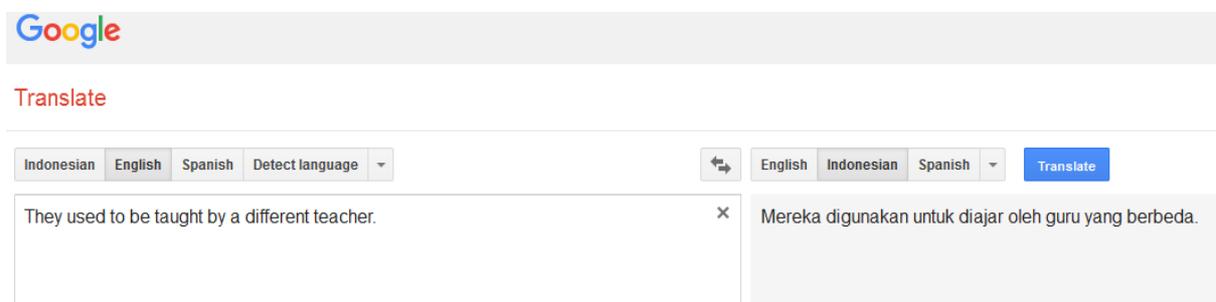


Figure 2: The translation of “They used to be taught by a different teacher”

The translation of “They used to be taught by different teachers” resulted in a global error. The error in this translation is caused by the failure of this machine to identify the phrase ‘used to’ correctly. The machine only sees the word ‘used’ per se so that in translation there is a confusion between the phrase "used to" and the word "used", which causes a significant deviation in meaning. The sentence “They used to be taught by different teachers” becomes “They are utilized to be taught by different teachers”, which is incomprehensible. Table 4 below shows the sentence pair:

Table 4: The sentence pair of They used to be taught by a different teacher.

	<i>Mereka</i>	<i>digunakan</i>	<i>untuk</i>	<i>diajar</i>	<i>oleh</i>	<i>seorang</i>	<i>guru</i>	<i>yang berbeda</i>
They								
used								
to be								
taught								
by								
a								
different								
teacher								

The sentence pair as seen in table 4 forms a diagonal line, represented by the light-colored boxes, moving from top left to bottom right. The diagonal line indicates the match of the language corpora. Yet, there are also dark colored boxes which interrupt the light diagonal line. Those boxes represent the wrong selection of words in the target language, and it is in this area that the intended meaning is disrupted.

The Use of Grammatical Structure: conditional tenses in English-Indonesian Translations

Another shortcoming is that Google Translate has not been able to transform all the English grammatical constructions into the target language coherently and accurately. Some grammatical constructions can be translated fairly well; while others cannot. One example of a well-translated sentence is a ‘conditional sentence’. This sentence is translated appropriately to the target language when sentence construction begins with IF. Please note that in English the prescribed sentence is divided into ‘open conditional’ and ‘subjunctive or ‘unreal conditional’. Open conditionals are open, meaning a presumed event can occur if the conditions are met; while the conditional subjunctive is the opposite. Open Conditionals always begins with IF, while ‘unreal conditionals’, besides beginning with IF can also start with WERE or HAD. Observe the following examples.

- (1) If I were you, I would not let her go. (which is similar to ‘Were I you, I would not let her go’).
- (2) If I had not worked in that institution during my study, I would not have been able to pay tuition. (which is equivalent to ‘Had I not worked in that institution during my study, I would not have been able to pay my tuition’)

The results of this study indicate that sentences beginning with IF can be translated easily; conversely *Google Translate* incorrectly translated all sentences beginning with WERE or HAD . able 5 demonstrates this tendency.

TABLE 5. The sentence pairings of “*Had he known that you were coming, he would not have left*”

	Apakah	Dia	Tahu		Bahwa	Anda	Akan	Datang	Ia	Tidak	Akan	meninggalkn
Had												
He												
Known												
That												
You												
Were												
Coming												
He												
Would												
not												
Have												
left												

Table 5 shows a diagonal trend that moves from the top left to the bottom right, a sign that sentence patterns move with a one-to-one relational pattern. But there is an inaccurate translation of the word ‘had’ which is simply to introduced the unreal conditional but recognized by the machine as an indicator of the perfect tense that serves to form interrogative sentence construction. Therefore, the translation is inaccurate. This is an example of a limitation of Google Translate related to the use of structure in the target language.

Another drawback is seen in the translation of sentences that contain ‘present participles’. Although the translation of a sentence containing the ‘present participle’ does not imply a global error in the target language, this inaccuracy of translation reduces the accuracy value of the machine translation. Consider the following examples:

- 09 Feeling hungry, he went into the kitchen and opened the fridge.
 10 Putting on his coat, he left the house

The sentences were translated as

- 11 *Merasa lapar, ia pergi ke dapur dan membuka kulkas .*
 12 *Memakai mantelnya, ia meninggalkan rumah*

The ‘present participle’, represented by ‘verb+ing’ in both sentences form an ‘adverb of reason’ (09) and an ‘adverb of time’ (10). Without the presence of ‘reason’ and ‘time period’ markers, these two sentences are still understandable, but the accuracy of the translation is reduced. Both sentences will be more appropriate if translated as follows:

- 13 *Karena merasa lapar, ia pergi ke dapur dan membuka kulkas.*
 14 *Setelah memakai mantelnya, ia meninggalkan rumah.*

Thus, the difficulty of transforming a complete grammatical construction of a source language into a target language, in this case, from English to Indonesian is one of the shortcomings of the Google Translate.

Conclusion and Suggestions for Future Research

As demonstrated in the discussion, the preliminary findings using Osbourne's methodology of analysis indicate that *Google Translate* is not only able to translate words and phrases. It has the capability to translate sentences with an accuracy level reaching 60.37; which is much higher than that of other Asian languages as reported by Patil and Davis (2014), which averaged 46 percent accuracy. Yet, this research also revealed that *Google Translate* struggled to translate sentences from English to Indonesian in instances where sentences contain words that have multiple functions and sentences which have different sense of language relating to point of view or perspective. As shown in my sample of sentences, *Google Translate* seems to frequently mix up the word 'that', which can have numerous different functions in a sentence. Google was unable to perform the translation until a human translator interfered and changed the perspective.

As I admitted earlier, this research is only a preliminary exploration of the accuracy of *Google Translate* when translating sentences from English to Indonesian; with a goal of shedding light onto the potential advantages and shortcomings of *Google Translate* when used to translate English sentences to Indonesian. Yet, it strongly indicates that despite its limitation, the machine is still able to translate sentences fairly well, with an accuracy of more than 60% based on Memsources criteria. This research is important in that it helps realize that judging the capability of machine translators through a lens of assumptions is ineffective. Users should utilize google translate with a clear understanding of its benefit and uses, as well as its drawbacks and areas of weakness.

This research itself has its limitation in that it only uses a limited source of data with limited coverage, which regard to difficulty experienced by Indonesian language learners at Inculs. It is thus challenging for future researchers to explore more on the accuracy of *Google Translate* translating Indonesian text to English using corpus data.

References

- Aitken and Balan. (2011). 'An Analysis of Google Translate Accuracy'. *Translation Journal*. Retrieved from <https://scholar.google.com/citations?>
- Barreiro, et al. (2014). *Linguistic Evaluation of Support Verb Construction by OpenLogos and Google Translate*. Retrieved from <http://www.cs.cmu.edu/~lingwang/papers/lrec2014-2.pdf>.
- Butler. (2010). *Machine Versus Human: Will Google Translate Replace Professional Translators?*. Retrieved from <http://mason.gmu.edu/~fbutler2/IT%20103-005%20Research%20Paper%20Butler.pdf>
- Farlin, S. (2015). *Semantics analysis in the translation of Indonesian abstract into English using Google Translate*. Retrieved from <https://repository.usd.ac.id/3782/1/081214124.pdf>.
- Graesser, Li, and Chai. (2014). 'Comparison of Google Translate with Human Translation'. *Proceedings of the Twenty Seventh International Florida Artificial Intelligence Research Society Conference*. Retrieved from <https://pdfs.semanticscholar.org/1187/d4bc0c83804c15cd6cc1b43670d27f5fe9b6.pdf>
- Grajales. (2015). *Statistics Behind Google Translate*. Retrieved from www.statisticviews.com/details/feature/806.081/The-statistics-behind-Google-Translate.html.
- Hardin & Picot. (1990) 'Unchangement de point de vue qui permet d'exprimer de manière différente une même phénomène'. *Translate: Initiation à la pratique de la traduction*, Bordas, Paris: Aubin Imprimeur, p. 21
- Koehn, P. Och, F.J. and Marcu, D. (2003) 'Statistical Phrase Based Translation'. In *Proceedings of the 2003 conference of the North American Chapter of the Association for Computational Linguistics on Human Technology*, Vol 1, 48-54. Edmonton, Canada: Association for computational Linguistics.
- Liputan 6. (2018). *Alasan Orang Indonesia Doyan Pakai Google Translate*. Retrieved from <https://inet.detik.com/cyberlife/d-3485785>
- Memsources (2016). *Data: Machine and Professional Human translations Identical in 5-20% cases*. Retrieved from <https://www.memsources.com/blog/2016/06/28/machine-vs-human-translation>
- Napitupulu. (2017). "Analyzing Indonesian-English Abstracts Translation in Views of Translation Errors by Google Translate". *International Journal of English Language and Linguistics Research* Vol.5, No.2, pp.15-23
- Ney, H. (1995). 'On the Probabilistic Interpretation of Neural Network Classifiers and Discriminative Training Criteria'. *IEEE Transactions on Pattern analysis and Machine Intelligence* 17 (2); 107-119

- Okpor. (2015). Machine Translation Approaches: Issues and Challenges. *IJCSI International Journal of Computer Science Issues*. Vol 11. Issue 5. No 2. September 2014
- Osborne, et al. (2006). 'Improved statistical machine translation using paraphrases', *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, p.17-24, June 04-09
- Osborne. (2010). Metrics for MT Evaluation: Evaluating Reordering". DOI: 10.1007/s10590-009-9066-5
- Papineni, et al. (2002). 'BLEU: A Method for Automatic Evaluation of Machine Translation'. In *Proceedings of the 40th Annual Meeting on association for Computational Linguistics*, 311-318, Philadelphia.
- Patil and Davies. (2014). *Use of Translate in Medical Communication: Evaluation of Accuracy*. Retrieved from <http://www.bmj.com/content/bmj/349/bmj.g7392.full.pdf>
- Sneddon, J. (1996). *Indonesian: A Comprehensive Grammar*. London : Routledge.