

## Equivalence Levels of Literary Corpus Translation Using a Freeware Analysis Toolkit

Frans Sayogie\*

Universitas Islam Negeri Syarif Hidayatullah Jakarta  
Jakarta, Indonesia  
[frans.sayogie@uinjkt.ac.id](mailto:frans.sayogie@uinjkt.ac.id)

Moh. Supardi

Universitas Islam Negeri Syarif Hidayatullah Jakarta  
Jakarta, Indonesia  
[moh.supardi@uinjkt.ac.id](mailto:moh.supardi@uinjkt.ac.id)

### **Abstract**

*Machine translation has the potential to make huge contributions to translation industries, but it seems, for now, that machine translation equivalence has led to a crucial point for literary translation by using machine translation because of the problem of the equivalence itself. This paper, therefore, aimed to see the equivalence degree of literary translation resulted by machine translation, i.e., freeware toolkit AntConc 3.5.0 software 2019. The data were collected from English-Indonesian J.K. Rowling's Harry Potter and The Order of The Phoenix novel by using the software to find the equivalent translation. The collected data were analyzed qualitatively based on the strategy of translation equivalent level proposed by Mona Baker (1992). The analyzed data revealed that the equivalent level of the software mostly occurred in a word level, above word level, and grammatical level. The software was likely difficult to find a textual level and a pragmatic level of translation equivalence because they required a context and still needed human involvement as part of a greater creative project of translation which were not done by the machine translation. After all, AntConc 3.5.0 as Computer Assisted Tool (CAT) brought a huge contribution to translation industry, helped to analyze large corpora particularly to find the degree of translation equivalence in word and above word level.*

*Keywords: CAT tool; equivalence level; literary translation.*

\*Corresponding Author

## Abstrak

Terjemahan mesin memiliki potensi untuk memberikan kontribusi besar terhadap industri terjemahan, namun demikian tampaknya, hingga saat ini, kesepadanan terjemahan sastra dengan menggunakan mesin terjemahan masih menjadi masalah yang krusial dikarenakan permasalahan kesepadanan itu sendiri. Oleh karena itu, paper ini bertujuan untuk mengetahui tingkat kesepadanan terjemahan sastra yang dihasilkan oleh terjemahan mesin, yaitu perangkat lunak freeware toolkit AntConc 3.5.0 2019. Data penelitian ini diambil dari J.K. Novel Rowling Harry Potter dan The Order of The Phoenix novel Inggris – Indonesia dengan menggunakan software Antconc 3.5.0 2019 guna menemukan kesepadanan terjemahan yang dihasilkan oleh software tersebut. Data penelitian dianalisis secara kualitatif dengan menggunakan strategi tingkat kesepadanan terjemahan yang diajukan oleh Mona Baker (1992). Hasil analisis data menunjukkan bahwa tingkat kesepadanan terjemahan software tersebut masih bisa bekerja dengan baik pada tataran kata (word level), di atas kata (above word level), dan tataran grammer (grammatical level). Sedangkan untuk tataran teks dan pragmatik, software tersebut tampaknya masih belum mampu mengidentifikasinya karena persoalan konteks yang masih membutuhkan keterlibatan manusia dalam mengidentifikasi hasil terjemahan yang lebih kreatif yang tidak dapat dilakukan oleh mesin terjemahan. Namun demikian, Antconc 3.5.0 sebagai Computer Assisted Tool (CAT) telah memberikan kontribusi yang sangat besar terhadap industri penerjemahan, dapat meringankan pencarian data korpus yang sangat besar, salah satunya adalah untuk mengidentifikasi tingkat kesepadanan terjemahan pada tingkat kata dan di atas kata.

*Kata kunci:* perangkat CAT, tingkat kesepadanan, terjemahan sastra.

## ملخص البحث

أدى البحث عن التكافؤ بين الترجمة البشرية والترجمة الآلية إلى الوجود المهمة كثيرا ما من مترجمين الأدبية بسبب غموض فكرة التكافؤ نفسها. ومن ناحية الأخرى تؤدي هذه الغموض والاختلافات في الترجمة الأدبية والنوع الأدب—إلى الدرجات المتفاوتة من التكافؤ. وهذه المظاهر تحت إلى أن تبحث على استكشاف مستوى التكافؤ في ترجمة مجموعة أدبية باستخدام تطبيق مجاني على الأعمال الأدبية من خلال مقارنتها بالمستوى المكافئ في الترجمة البشرية. أسلوب بحث نوعي مع تصميم تحليل محتوى الترجمة باستخدام “مجموعة أدوات تحليل مجانية”. تركز الترجمة الأدبية من خلال مجموعة مرجعية على الكلمات والجمعيات والعبارات والجملة التي يقوم بها المترجم بمساعدة البرنامج المجاني المسمى ب”AntConc 3.5.0 2019” لتحليل بيانات البحث. وحاصل نتيجة البحث يدل على أن الترجمة الأدبية لا تزال نشاطاً متعاضدا الذي يحتاج إلى إشراك البشر. وهي أيضا أن مستوى معادلة الترجمة وجد على مستوى معادلة الكلمات، فوق الكلمة، والمستويات النحوية، والنصية، والبراغماتية. ومع ذلك، فقد خلاص البحث إلى أن معظم أوجه التشابه تحدث على مستوى الكلمة وفوق مستوى الكلمة.

المصطلحات الأساسية: جهاز CAT ؛ مجموعة أدوات تحليل مجانية؛ مستوى المقارنة؛ الترجمة الأدبية .

## INTRODUCTION

Many translation studies using corpora had been developed over the years that might serve as a remedy, evaluation, and enhancement of the quality of translations (De Sutter, 2017 and Arcan, 2018). Digital corpora as computer-assisted tools (CAT) to develop translation quality has been used to build an understanding of the equivalence level of naturalness (Frankenberg, 2015; Floranti & Mubarok, 2020). Regarding the notion, a literary translation resulted from non-human has been reacted differently because of the quality of translation equivalences. If there is no equivalence in literary translation works, the untranslatability issues arise (Sayogie & Supardi, 2019). Literary translation is a worth mentioning that the *Divan of Hafez* is replete with puns, homonyms, and wordplays, which although being interesting for the Persian readers, seems to bore, incomprehensible and intolerable for English readers. Some of them find literary translation interesting, while others find literary translation annoying (Hakemi, p. 376). Regarding the notions, the present research endeavors to seek the equivalence problems of literary translation using a freeware toolkit AnConc 3.5.0 introduced in Anthony Laurence website.

The technology of machine translation (MT) has been both significant and bad for translators. Translators have been helped by technological peripherals such as internet search engines and online dictionaries and encyclopedias. Some translators are also satisfied with translation tool technology and find that it has made their work more approachable and real. CAT tools can reduce the time translators spend on working repetitive tasks, but some literary translators discard CAT tools for doing their work (Ehrensberger-Dow, 2017). Besides, translation technology can decrease qualified translator independence and threaten the livelihood of professional translators (Taivalkoski-Shilov, 2019).

Literary translation, somehow, has become one of the most interesting topics in the subject of translation research in recent years. The changing development of literary translation required translation techniques and changes in the reader's cultural expectations and aesthetic concepts (Robinson, 1997). One of the most critical problems comes from the translator itself, such as reliability, involvement in the profession, ethics, good intelligence, and memory, also the capability of creating the link between the text and the intended readers (Hermans, 2007). Professional translators, therefore, are expected to loosen up in examining the problem of words, phrases, syntactic or grammatical structures, textual and pragmatic aspects, and a cultural assumption, using analytical awareness of the problem and its possible answer. As a result, it is essential to understand how to be an excellent translator and acknowledge every aspect to deal with translation problems.

Translation still struggles with recognition and undervaluation even after its growth as a profession in the last century (Flanagan, 2016; Tyulenev, 2015). Most of the translators, in several cases, still entrust in the processes for translation technologies (Cadwell, 2017; Doherty & Dorothy, 2014). They protest technology because of the necessity to defend their jobs and appeal to the defense of quality of the translation itself. This demeanor is neither an alternative approach nor something limiting professional translator works.

The development of translation technologies is still growing, and the quality of translation by using a freeware tool kit has enhanced considerably. Research has also supported rapid productivity through the technology of software translation (Green, 2013; Koponen, 2016). Although the quality of machine translation is increasing today, the raw results of machine translation are considered not very useful for direct use. This is because some output still contains some errors of equivalence in the Target Language (henceforth TL) (Lotz & Van Rensburg, 2016). It is assumed that the translation work is an activity with human involvement. The problems of equivalence and untranslatability become real, and translators have strived to find results to non-equivalence and untranslatability because

regardless of the strategies of translation used the definitive target preserves to signify the TL (Anh, 2018). Anything short of that can be measured as an unsuccessful translation.

A great deal of theories has been offered by many scholars (for instance, Newmark, 1981, Vinay and Darbelnet 1995, Nida 1964, Toury, 1995, Munday, 2008) to overcome the problems of translation. The words order of literary translation, such as the style, will cause the change between a zippy, extremely readable translation and a fixed, and artificial altering that deprives the uniqueness of its aesthetic core, even in its very soul (Bassnett, 2014). Besides this, literary translators must concern with the flexibility, style, tone ingeniousness, the source language (henceforth SL) culture knowledge, the capability to grasp the meaning from ambiguity, capable of listening sonority, and humility.

This study may seek to find whether the changes are because of the differences in the genre, translation techniques, a student's degree (amateur or professional), or any other factors. Besides this, different "basic linguistic features" should be included in the analysis, such as the lexica-grammatical indicators mentioned in (Evert & Neumann, 2017). Examining translation research through corpora was first introduced by Baker in 1992. This new partnership corpus linguistics may serve the method in conveying empirical studies while the theory of translation may describe the parts of inquiry and elaborate operational hypotheses. The equivalence levels of translation were encountered in the equivalence levels of word, above word, grammatical, textual, and pragmatic through the literary text translated by CAT AntPconnc 3.5.8.

## METHOD

The study of examining translation and translating through linguistic corpora was firstly taken by Baker in 1993. Such a study had influenced the researcher to adapt similar research through the literary translation corpora by using CAT AntPconnc 3.5.8 as the data collection of this research. The corpora offer the method for conducting an examination while translation theory would recognize the areas of review and run operational hypotheses. The partnership between linguistic corpora and translation has developed a simple character with a specific value, Corpus-Based Translation Studies (CTS). Corpus based translation studies is therefore a method or an approach used to investigate translation phenomena whose definition can be safely gleaned from the corpus linguistics definitions offered by Baker (2010) and McEnery and Hardie (2012).

The concept of equivalence levels by Mona Baker is used to analyze the corpus-based approach to literary translation equivalence (Baker, 2011). The source of data is "J.K. Rowling's Harry Potter" and "The Order of The Phoenix" Published by Arthura Levine Books, an imprint of Scholastic Inc 2003 based on a translator training using computer-assisted tool AntPConn 3.5.8 2019, and its Indonesian translation as the corpus of the study. All the English–Indonesian translation was only analyzed randomly.

The data were collected through a reference corpus focusing on the most common words, collocations, phrases and sentences that were numbered based on the segment of translation. They were excluded from the list of frequent words in the corpora of AntConc software, including a reference corpus for translation purposes to get an overview of keywords in the corpus. These keywords were extracted by the AntConc software tool through a comparison of the two corpora, English and Indonesian texts. The compilation of the corpora is considered a onetime comparison for the extraction of the corpus. The corpus should be proportional to the size of the corresponding corpus. Through reading, marking, classifying, comparing between SL and TL in the matter of non-equivalence and close equivalence in the word level, above word, grammatical, textual, and pragmatic equivalence using AntConc Version 3.5.8 2019, the concordance lines for a parallel corpus of data would

be found. Here, linguistic corpora can contain texts from different modes of language elements.

The data were analyzed qualitatively by using the translation strategy proposed by Mona Baker (1992) to analyze equivalence level of words, above words, grammatical, textual, and pragmatic. The data resulted by the AntConc software were used to the retrieval of the corpus as the linguistic data and to track linguistic equivalence in the corpora, such as the rank, frequency, and words or sentences.

## FINDINGS AND DISCUSSION

### Equivalence in Word Level

In analyzing a freeware tool kit AntConc relating to translation equivalence, four different levels of the notion of equivalence described in Baker's were used, i.e., words, grammatical, textual, and pragmatic equivalence related with the translation procedure. The word equivalence can occur at a word level and above word level when reproducing text from SL into TL. Equivalence at a word level is the primary level, which is taken into some considerations by translators. Here, translators should know several aspects when rendering a single word such as "fall" in "I love cool, crisp fall weather" or "Don't fall on your way to the gym." For grammatical equivalence, there are many grammatical rules across languages that may change the meaning of SL text in TL text. These changes may bring the translator either to omit or to add material in the TL text. Thus, textual equivalence is essential in a translation that will help the translator create a cohesive and coherent text in a specific context. The last one is the pragmatic equivalence where translator must consider implicatures and cultural differences among languages during translation process (Baker, 2011).

Figure 1 provides translation resulted from a freeware AntConc. It is considered that literary corpus translation applies to get a sense of how a freeware use differs slightly from human translation.

Word Types: 11255			Word Tokens: 113564			Word Types: 11255			Word Tokens: 113564			Word Types: 11255			Word Tokens: 113564		
Rank	Freq	Word	Rank	Freq	Word	Rank	Freq	Word	Rank	Freq	Word	Rank	Freq	Word	Rank	Freq	Word
1214	12	berbalik	1371	11	tegas	147	112	can	1678	8	fingers	148	112	now	1679	8	forward
1215	12	berbaring	1372	11	terbangun	149	111	did	1680	8	friends	150	111	off	1681	8	furious
1216	12	berubah	1373	11	thinks	151	110	when	1682	8	grin	152	109	again	1683	8	gryffindor
1217	12	blood	1374	11	try	153	109	right	1684	8	hak	154	109	time	1685	8	halo
1218	12	bunga	1375	11	twins	155	107	begini	1686	8	heading	156	107	going	1687	8	hebat
1219	12	buru	1376	11	ungu	157	105	kreacher	1688	8	held	158	105	looked	1689	8	hello
1220	12	business	1377	11	urusan	159	103	here	1690	8	herself	1221	12	corner			
1221	12	corner	1378	11	usaha							1222	12	cover			
1222	12	cover	1379	11	used							1223	12	cuma			
1223	12	cuma	1380	11	waktunya							1224	12	departemen			
1224	12	departemen	1381	11	wardrobe							1225	12	dilakukannya			
1225	12	dilakukannya	1382	11	watched							1226	12	dunq			
1226	12	dunq	1383	11	whatever												

Figure 1. Equivalence in a Word Level

SL: "Not our business what he does at school, eh?"

TL: "Bukan urusan kami apa yang dia perbuat di sekolah, eh?"

SL: "Nor does it have the right to confiscate wands until charges have been successfully proven; again, as I reminded you on the night of the second of August."

TL: "Juga tidak mempunyai hak untuk menyita tongkat sihir hingga tuntutan telah dibuktikan dengan sukses; sekali lagi, seperti yang kuingatkan kepadamu pada malam dua Agustus."

The word 'business' in Figure 1 which is translated into *urusan* can be considered as equivalence translation at the word level. The word 'business' was identified 12 in frequencies and 1220 in the word rank. Even though the word "business" in SL was slightly lower than the word *urusan* in TL was 1377 in the word rank, but it is lower one level (11) in the frequencies. The translation of the word 'business' into *urusan* is considered close equivalence in the word level because of its naturalness and acceptability. The equivalence word can be understood simply by Indonesian readers. The concept of word "business" in SL has the same concept in TL and based on the lexical meaning that the word 'business' is commonly expressed as *urusan* in Indonesian culture. The lexical unit of the TL word has a specific cultural value in a particular linguistic system. The TL word has identical propositional meaning as the SL word, and it has an identical meaning. It is important to find a translation in a specified context. If the TL equivalent is neutral in comparison with the SL item, the translator may find the word which is acceptable according to the target readers' culture. The strategy used by the translator can be categorized as a communicative translation or dynamic translation (Xiyao, 2015).

For problems of non-equivalence at a word level that occurred using this tool, translators must know the cultural differences between languages when the output of translation is produced. Therefore, translators should have a good background of knowledge of both SL and TL (Ezzati, 2016).

The translation of the word 'right' in the SL is considered equivalence if it is translated into *hak* because it is familiar for Indonesian readers. The word 'right' which was translated into *hak* can be considered as equivalence translation at the word level. The word 'right' was identified 109 in frequencies and 153 in the word rank. The word 'right' in SL was much lower than the word *hak* in target language for 1684 in the word rank, but in frequencies, the word 'right' is much higher (109) compared to the target language (8). The translation of the word 'right' is considered or categorized as translation using semantic translation or formal translation. The translation is considered equivalence in the word level because of the natural effect for the readers, Indonesian readers. The meaning of the sentence has been transferred literally but can still be considered equal by the target language culture. This translation can be categorized as close equivalence or adequate translation because the meaning of the sentence can be grasped easily by the target language readers. The translation of words in the source text gives the same effect in translation. The translation may find a different way or word choice for different translators that match the context of the word.

### **Equivalence in Above Word Level**

In figure 2, the translation equivalence will move from word level to the above word level, i.e., phrases, collocation or idioms in which that collocations and other multi-word units such as two-verbs, and idioms are principally problematic. The difference will be more significant if they are technical terms in scientific texts, since the lexical unit is often longer than one word, as cited by Anthony (2005) (Bowker & Pearson, 2002). Amazingly, phrasal words, collocations have perceived brief consideration in most CALL programs (Anthony, 2005). Collocation is a notion at the syntagmatic level and is not directly related to the original meaning of words. A collocation establishes the connection between the terms in unfamiliar words. Collocation is very significant to comprehend language learning concepts. In TL, translators sometimes use inappropriate collocations since English words are not linked in collocations in the non-native speakers' memory (Shraideh & Mahadin, 2015).

Word Types: 11255			Word Tokens: 113564			Word Types: 11255			Word Tokens: 113564		
Rank	Freq	Word	Rank	Freq	Word	Rank	Freq	Word	Rank	Freq	Word
4876	2	glowing	1033	15	menyadari						
4877	2	gobstones	1034	15	menyuruh						
4878	2	godaan	1035	15	met						
4879	2	goldstein	1036	15	minggu						
4880	2	gonggongan	1037	15	pulang						
4881	2	goodness	1038	15	saling						
4882	2	goodnight	1039	15	selamat						
4883	2	gordon	1040	15	seorangpun						
4884	2	goshawk	1041	15	short						
4885	2	grabbing	1042	15	should						
4886	2	grasp	1043	15	started						
4887	2	greenland	1044	15	stepped						
4888	2	greeted	1045	15	steps						

Figure 2. Equivalence in Above Word Level

SL: "Goodnight"

TL: "Selamat malam"

The translation of the phrase in Figure 2 is considered non-equivalence in the above word level because of the phrase or idiom 'good night' from the source language is translated into *selamat malam*. Here, the words 'good night' was identified 2 in frequency and 4882 in the word rank. The word 'good night' in SL was much lower than the word *selamat malam* in target language for 1039 in the word rank but in frequencies, the word *selamat* is higher (15) compared to the source language (2). The words 'good night' in SL could only be identified by the AntConc software as *selamat malam* even though in the original text before they are input into the toolkit were translated into *selamat malam*. From this case, it implies that AntConc software cannot generate the equivalence translation as the original text. Such translation is considered not equivalence which is not acceptable in TL culture and the strategy used by the translator is categorized into translating idiom into the non-idiom in TL. For Indonesian readers, the word 'goodnight' should be collocation or sound natural if it is translated into *selamat malam*. The translator seems to consider the target readers' culture. So, the translation of 'good night' resulted by the AntConc toolkit is considered non-equivalence because it gives unnatural effect for the target readers. The SL concept expresses the problematic concept in TL culture because of the lack of the word *malam*. The translation is considered non-equivalence because it is unacceptable by target readers. This problem, however, was reasonable since the AntConc software is more effective to use in students' writing task to identify the style, word rank or frequency, collocation, and to find the lists of the words. Another example of above word level translation is available as follows.

SL: "And all of a sudden, for the very first time in his life, Harry fully appreciated that Aunt Petunia was his mother's sister"

TL: "Dan tiba-tiba, untuk pertama kalinya dalam hidupnya, Harry benar-benar menyadari bahwa Bibi Petunia adalah kakak ibunya"

The translation is considered close equivalence because it gives the same effect for the TL readers. Here, the equivalence level of the translation was displayed in the concordance hit, which is focused on the level of the above word 'all of a sudden.' In this example, the concordance hit showed the different result in a number, which is very remote different between the SL and the TL. Based on figure 4, the words 'all of a sudden' was encountered in three different sentences. Among the three different usages, there are similar cases in number, i.e., 960, 961, and 962 according to a high index of concordance hits. Surprisingly, the number of frequencies of the source language is much larger than the number of rates in TL, which range only 1,2, and 3 showed by the concordance hit.

This might suggest evidence that the differences result between the SL and TL show different levels of equivalence based on the number of collocation or above words level among them. Even though they are different, the result of translation is considered equivalence. According to the concept of equivalence translation proposed by Mona Baker, such translation can be categorized as equivalence in the above word level because the phrase or idiom 'and all of a sudden' cannot be translated literally. Instead, the translator translates it into a different form, but is still acceptable to the target readers. The SL brings the same concept to the TL using the strategy of translation idioms by using non-idioms in TL. There are two main types of concordances: first, these can build an index that has a function for subsequent search operations, and second, those that act directly on the raw text (Hockey, 2001; Antony, 2005). The first of these has the benefit that they can operate on large corpora. Besides, they are used very often to switch or change the target corpus for a particular need.

To perform all processing on the raw data, files, and storing results in inactive memory, AntConc fits into the second type. That is why using it with small specialized corpora is restricted. However, one of the significant tendencies in corpus linguistics over the past few years, as McEnery and Wilson note (McEnery & Wilson, 2001), is the increased substance in very small, highly specialized corpora. Ghadessy exemplified small corpora can be used for many different purposes and are especially useful if it is used to teach technical writing (Noguchi, 2004).

### Equivalence in Grammatical Level

For the following figure we can see equivalence in grammatical level. Baker (2011) uses grammatical equivalence when denoting to the variety of grammatical classes throughout languages. She states grammatical guidelines may differ throughout languages, and this may pose some cases to find a correspondence in the TL. She claims that different grammatical structures in the SL and TL may change significant meanings in the source language text (henceforth SLT) when rendering in the target language text (henceforth TLT). Baker also states that the non-equivalence frame of grammatical level in a language has a unique configuration that is not adequate in another language. Besides, people of target language use entirely different tense and aspect of sentences and words to get the same meaning. These grammatical categories are easy to comprehend for the speakers of a similar language. These categories or elements would be nonsense to the interlocutors of the other languages. The elements of language grammatical create that the SL has some substances at word level that is entirely different from the corresponding objects in the TL.

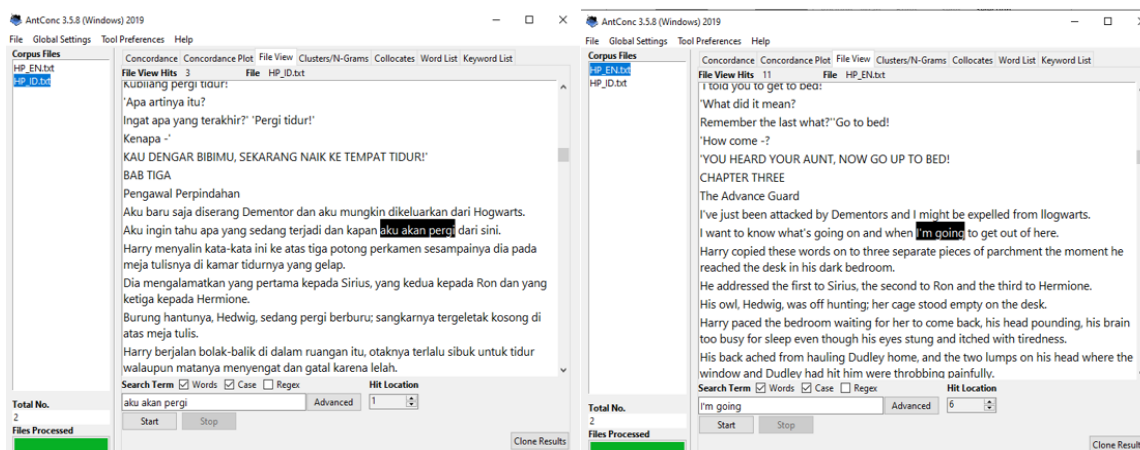


Figure 3. Equivalence in Grammatical Level



SL: "I'm going"  
 TL: "Aku pergi"

In Figure 3, the translator translates the sentence 'I'm going' into *aku pergi* which is categorized into a grammatical level in the tense-aspect. The translation is considered equivalence in the grammatical level because it can be accepted by TL readers. The meaning of the sentence still has the same effect for target readers. The translator translates the sentence by using a different tense from the source language tense. The translator's decision to translate different tense from the source language can be considered as communicative translation. If the source language structure or tense is translated literally, the translation will be awkward or clumsy.

The present continuous tense in SL is considered not typical because of the different norms and structures. There may be a TL word that has the same correspondent meaning as the SL word, but it may usually be unlike in expressive meaning. The difference may be considerable, or it may be implied, but it is significant enough to present a translation problem in a specified context. If the structure of TL is less equivalent to target language readers, they are supposed to be altered based on target readers' norms. Differences in expressive meaning are usually more challenging to handle when the TL equivalent is not available in TL structure. That is why the translators often face difficulties and make some adjustments by changing the structure or tense. The strategies used by the translator should adjust the grammatical rules according to the TL.

When a translator deal with a language that has several dissimilarities into a language with no aspect type, the common problem may emerge. The above example of strategies used by the translator can be categorized into translation by using communicative or dynamic translation. Regarding the problem of collocation, most corpus analysis programs suggest translators the ability to perceive the collocates of a search term as shown in the figure 3 where the frequency of the most common words to the left or right of the search term are indicated. The recent version of AntConc, therefore, offers no implementation of this feature, since learners often find such tables difficult to understand. Several programs also provide detailed statistics concerning the search results and corpus (L, 2005). Another translation in the grammatical level is found in the sentence below.

SL: "Before - anyone - sees!"  
 TL: "Sebelum - dilihat – orang lain!"

The translation of the word 'before – anyone – sees' into *sebelum – dilihat – orang lain* is categorized into a grammatical level in the voice aspect. The active sentence in SL is translated into passive in TL. The translation is considered equivalence in the grammatical level because it can be accepted by TL readers. The meaning of the sentence still has the same effect or meaning. The translator translates the sentence by using a different structure with the SL structure. The translator's decision to translate different structures from the SL can be considered as communicative translation. If the SL structure is translated literally, the translation will be awkward or clumsy.

The active voice in SL is rendered into passive voice because of the different norms and structures between English and Indonesian. It is usually different in expressive meaning, since there may be a TL word that has the same correspondent meaning as the SL word. If the structure of TL is less equivalent to target language readers, they are supposed to be altered based on target readers' norms. When the TL equivalent is not available in TL structure, differences in expressive meaning are usually more difficult to render. That is why

the translators often face difficulties and make some adjustments by changing the information. The grammatical rules in TL can cause certain strategies used by the translator. The strategies used by the translator in the example above can be categorized into translation by using communicative or dynamic translation.

Baker (2011) defines the textual equivalence as coherence and cohesive elements that embrace the text between the SL and TL. The textual equivalence will be accomplished when the cohesive devices are strictly used when translating the text.

### **Equivalence in Textual Level**

*SL: "He cast around for a topic that didn't involve his headmaster because the very thought of Dumbledore made Harry's insides burn with anger again"*

*TL: "Dia memandang berkeliling mencari topik yang tidak melibatkan kepala sekolahnya, karena memikirkan Dumbledore saja membuat tubuh bagian dalam Harry terbakar oleh amarah lagi"*

In the above example, the source language text is united by several cohesive words that build a sentence unity. The words that are underlined were considered as the cohesive word that unites the sentence. The translator translates the sentence "He cast around for a topic that didn't involve his headmaster because the very thought of Dumbledore made Harry's insides burn with anger again" into *Dia memandang berkeliling mencari topik yang tidak melibatkan kepala sekolahnya, karena memikirkan Dumbledore saja membuat tubuh bagian dalam Harry terbakar oleh amarah lagi*. The translator preserved the cohesive words in SL but in different forms. Such translation is considered equivalence in the textual level due to the many cohesive words that can unite the words into a sentence.

The textual level in Mona Baker's concept is the same as the sentence. The translation is considered equivalence in the textual level because of the cohesive words in SL are translated literally but still sounds for target readers. There must have been a TL word that has the same correspondent meaning as the SL word, but it may have a different expressive meaning. It is significant enough to pose a translation problem in a specified context, but the difference may be substantial, or it may be implied. If the structure of TL is less equivalent to target language readers, they are supposed to alter based on target readers' norms. According to Mona Baker, cohesion is objective and based on the principle of automatic recognition, in contrast to coherence that is more dependent on the reader's attitude. The counterparts of the text are connected by words and expressions cohesion.

The surface expression of coherence relations, as Baker calls it a cohesion device, make conceptual relations explicit. When the TL equivalent is not available in TL structure, differences in expressive meaning are usually more difficult to understand, since not every language has the same linguistics structure. That is why the translators often make some adjustments and face difficulties in changing the information. If it is so, the strategies used by the translator will depend on the textual cohesiveness in TL.

### **Equivalence in Pragmatic level**

*SL: "Dudley Demented"*

*TL: "Dudley Diserang Dementor"*

*SL: "Dudley's terrified voice broke in Harry's ear"*

*TL: "Suara Dudley yang ketakutan sampai ke telinga Harry"*

SL: *"I'm a Squib, as Mundungus knows full well, so how on earth was I supposed to help you fight off Dementors?"*

TL: *"Aku Squib, seperti yang diketahui Mundungus dengan baik, jadi bagaimana mungkin aku dapat menolongmu menghadapi Dementor?"*

SL: *"But I was getting rid of Dementors, I had to use magic - they're going to be more worried about what Dementors were doing floating around Wisteria Walk, surely?"*

TL: *"Tapi aku tadi mengenyahkan Dementor, aku harus menggunakan sihir - mereka pasti lebih khawatir tentang apa yang dilakukan Dementor melayang-layang di sekitar Wisteria Walk?"*

SL: *"It was a couple of Dementors!"*

TL: *"Tapi sepasang Dementor!"*

SL: *"And what the ruddy hell are Dementors?"*

TL: *"Dan apa itu Dementor?"*

The translation of 'Dudley demented' which is translated into 'Dudley diserang dementor' is considered as pragmatic equivalence since the entire story is talking about "how Dementor hunts dudely". The word "dementor" itself according to the story is still unclear for what exactly Dudley saw. We can speculate and also go off of what Rowling has said, that a Dementor near Dudley forced him to see his greatest fear, which was him exactly as he was. If that image of himself is Dudley's worst fear, then that the Dementors themselves bring out the worst fear in a person, even if they don't even realize what the fear is (or both). Such translation is considered equivalence in the pragmatic level because of the same implication aspects.

Baker states pragmatics as the study of language in use. The system of linguistic does not result in the study of meaning. However, it is but as carried and falsified by interlocutors in a communicative situation. Pragmatic equivalence has something to do with the reader's attitude and theoretical issues, which emphasizes the value of the text, including emotional connotations and its culture. The translation can be categorized into a pragmatic level in the smallest unit by using a translation strategy of paraphrase using unrelated words in different aspects of implication. The expressive meaning of the translated word or the level of expressivity is, sometimes, weakened or intensified by the target word used. Instead, the translator tries to preserve the similar expressivity level, which can be seen from the context. Translators have to evade pragmatic insufficiency in translation. Translating the clear meaning is not adequate to gain good pragmatically translations. Translators are obligatory to use the deep meaning of the original and other strategies in translation such as modifications, additions, omissions and deletions. Baker exemplifies that the output that is probable to disrupt the TL reader's expectations must be thoughtfully observed to evade transmission of the incorrect implicatures to produce sense. These adjustments are taken together under modification category. The modification strategies play between two different extremes, i.e., preservation or adaptation. Adaptation is a term which means changes made in the TLT when the context mentioned in the SLT does not occur in the TL culture. The purpose of adaptation is to keep the original text function (Baker, 2011).

After examining literature translation output using a translation tool, some diverse classifications definitely point to some errors that must be corrected with an additional context or cultural information. Here are the proposed classes in details. The level of words or phrases in the SL are translated into the TL with the different meanings of words or phrases, and this output of translation is misleading for the reader when reading the output in the TL. The reader should discover the original meaning in the TL sentence. In most cases, this is unclear words or phrases in the SL that is supposed to be replaced with the

correct sense of meanings. The minor meaning errors of translated words or phrases convey the SL original meanings. The equivalence of words or phrases that occurred in the output has a very different meaning or is not suitable for the specified context. The consistency for a term translation error explicitly addresses translation reliability for words and phrases that should have a different translation throughout the text. With idiomatic translation, the translator must attempt to recreate meanings of a similar equivalence meaning of the idioms from the SL (Floranti, 2020). Then, the output comprises syntactic or morphological mistakes that distort meanings in the TL (Matusov, 2019).

The data evaluation above is that literary translation works need an outstanding appreciation that is not possible for machine translation (MT). MT is more appropriate for non-literary translation to be used. Even for non-literary texts, it is not possible to rely on machine translation entirely. Language is very natural, communicative, active and original while the machine translation with encoding-decoding mechanism of computer process stresses accurateness and firmness and cannot tolerate vagueness. It is difficult to gain a perfect match between MT and natural language use (Zheng, 2015).

In this evaluation, the text must have been edited by a translator for making the non-equivalence translations in TL segments to be acceptable, however it does not attain the same quality level of equivalence between SL and TL. Both machine and human assessments show that specifically personalized systems using a literary corpus perform much better than general-purpose commercial systems. The quality levels got with the customized methods are good enough to handle the MT system in exact conditions (Hadley et al., 2019). Equivalence evaluation is a good peripheral for translators to get the high-ranked quality of the product of translation for readers. Without equivalence evaluation, to get an excellent quality of translation is not possible. This evaluation can offer a technique to measure or succeed a process correctly and also is a way for gaining readers' satisfaction (Gharedaghi et al., 2019).

## CONCLUSION

A freeware tool kit is a method for comparing and analyzing corpus linguistics or literary translation analysis in terms of translation equivalence process. Theoretically, this process would be the most important method of analysis in corpus linguistics, because it is a comprehensive way as it emphasizes on the text/corpora and selects the most frequent words for the analysis.

Even though the AntConc software as a freeware tool kit is not as useful as when it is used in teaching writing technique, because of its primary function is to keep a large corpus; it can also help translators' activities to compare a lot of corpus at one time. As a freeware toolkit, the AntConc software can also analyze parallel texts or sub-parallel texts. This software can compare translation meaning, especially in the word level or above word level. There are many complicated factors that are culturally perplexed to get equivalence translation of a novel the TL readers. Because of differences in form and the usage of loan words in the SL text, the SL word is semantically complex, and the TL lacks a specific term, SL concept is lexicalized in TL, the absence of culture-specific concepts, the SL and TL make different in meaning, the TL absences a superordinate, differences in physical-relational perspective.

Despite its accuracy, meaning cannot be obtained adequately in analyzing grammatical, textual, and let alone in the pragmatic level. Still, it may help to explain corpus linguistics or other text genres, such as literary text in the first steps. Therefore, this software is good to focus on the text or corpora and to select the most frequent words to be analyzed. To analyze literary translation, a translator cannot rely heavily on the result to a literary translation has

more than just language structure, i.e., cognitive aspects, social, cultural, and so on. Inevitably, a literary translation analysis needs the context of text or corpora rather than just the linguistic elements. The words meanings in a text are usually different from the hidden meaning that keywords may influence to extend in the other aspects. Thus, using multi-approach, bottom-up, or top-down, may help to make the analysis more accurate. But for this literary translation work, we assume the tool is not appropriate for literary translation.

## REFERENCES

- Anh, T. V. (2018). *Strategies For Non-Equivalence At Word Level In Literary Translation – A Case Study*. 11.
- Arcan, M. (2018). *A comparison of statistical and neural machine translation for slovene, serbian and croatian*. 3–10.
- Baker, M. (2011). *In other words: A coursebook on translation* (2nd ed). Routledge.
- Bowker, L., & Pearson, L. (2002). Bowker, L., & Pearson, L. (2002). Working with specialized language: A practical guide to using corpora. Routledge. *Journal*.
- Cadwell, P. (2017). *Resistance and accommodation: Factors for the (non-) adoption of machine translation among professional translators*. <https://doi.org/doi/full/10.1080/0907676X.2017.1337210>
- De Sutter, G. et. al. (2017). *Towards a corpus-based, statistical approach to translation quality: Measuring and visualizing linguistic deviance in student translation*. 16, 25–39.
- Doherty, S., & Dorothy, K. (2014). *The design and evaluation of a statistical machine translation syllabus for translation students*. 8(2), 295–315.
- Ehrensberger-Dow, M. (2017). An Ergonomic Perspective of Translation. In *The Handbook of Translation and Cognition* (pp. 332–349). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119241485.ch18>
- Evert, S., & Neumann, S. (2017). *The impact of translation direction on characteristics of translated texts: A multivariate analysis for English and German*. In G. De Sutter, I. Delaere, & M.-A. Lefer (Eds.), *Empirical Translation Studies: New theoretical and methodological traditions*. 47–80.
- Ezzati, A. (2016). Non-Equivalence at Grammatical and Word Level and the Strategies to Deal with: A Case Study of English Translation into Persian. *International Journal of Language and Linguistics*, 3(3), 7.
- Flanagan, M. (2016). Cause for concern? Attitudes towards translation crowdsourcing in professional translators' blogs. *The Journal of Specialised Translation*. *The Journal of Specialised Translation*, 25, 149–173.
- Floranti, A. D. (2020). Indonesia–English Translation of Idiomatic Expressions in The Novel This Earth of Mankind. *Buletin Alt-Turas*, 26(2), 14.
- Frankenberg, G. A. (2015). Training translators to use corpora hands-on: Challenges and reactions by a group of 13 students at a UK university. *Journal Corpora*, 210(3), 351–380.

- Gharedaghi, M., Eslamieh, R., & Shahidi, H. R. (2019). Content Equivalence Analysis of Health News Translation: A Bakerian Approach. *Theory and Practice in Language Studies*, 9(9), 1198. <https://doi.org/10.17507/tpls.0909.17>
- Green, S. (2013). *The efficacy of human post-editing for language translation*.
- Hadley, J., Popovic, M., Afli, H., & Way, A. (2019). *Proceedings of the 1st Workshop on Qualities of Literary Machine Translation*.
- Hakemi, B. G. (n.d.). *The possibilities and limitations of literary translation: A review of J. Payne's and Henri Clarke's Translations of Ghazalyat of Hafez*. 13.
- Hermans, T. (2007). *The conference of the tongues*. St Jerome.
- Hockey, S. (2001). "Concordance programs for corpus linguistics" in *corpus linguistics in north america: Selections from the 1999 symposium*. 76–97.
- Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation*, 25, 131–148.
- L, A. (2005). AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. *Professional Communication Conference*, 729–737.
- Lotz, S., & Van Rensburg, A. (2016). Omission And Other Sins: Tracking The Quality Of Online Machine Translation Output Over Four Years. *Stellenbosch Papers In Linguistics*, 46(0). <https://doi.org/10.5774/46-0-223>
- Matusov, E. (2019). The Challenges of Using Neural Machine Translation for Literature. *Proceedings of the Qualities of Literary Machine Translation*, 10–19. <https://www.aclweb.org/anthology/W19-7302>
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics. An introduction*. (Second edition). Edinburgh University Press.
- Noguchi, J. (2004). *A genre analysis and minicorpora approach to support professional writing by nonnative English speakers*. 11, 101–110.
- Robinson, D. (1997). *Becoming a translator: An introduction to the theory and practice of translation*. Routledge Group.
- Sayogie, F., & Supardi, M. (2019). Equivalence and Untranslatability in English Translations of UUD Negara Republik Indonesia 1945. *Proceedings of the 2nd Internasional Conference on Culture and Language in Southeast Asia (ICCLAS 2018)*. Proceedings of the 2nd Internasional Conference on Culture and Language in Southeast Asia (ICCLAS 2018), Tangerang Selatan, Indonesia. <https://doi.org/10.2991/icclas-18.2019.2>
- Shraideh, K. W., & Mahadin, R. S. (2015). Difficulties and Strategies in Translating Collocations in BBC Political Texts. *Arab World English Journal*, 37.
- Taivalkoski-Shilov, K. (2019). Ethical issues regarding machine(-assisted) translation of literary texts. *Perspectives*, 27(5), 689–703. <https://doi.org/10.1080/0907676X.2018.1520907>
- Tyulenev, S. (2015). *Toward theorizing translation as an occupation*. 2(1), 15–29.

Xiyao, H. (2015). *How to Achieve Equivalence, the Eternal Issue in Translation Studies: A Review of In Other Words: A Coursebook on Translation*. 4.

Zheng, H. (2015). A Case Study of Machine Translation: Problems and Suggestions. *International Journal of English Linguistics*, 5(2), p92.  
<https://doi.org/10.5539/ijel.v5n2p92>



© 2021 by Frans Sayogie, Moh. Supardi

This work is an open access article distributed under the terms and conditions of the Creative Commons Attribution-Share Alike 4.0 International License (CC BY SA)

Frans Sayogie, Moh. Supardi

*Equivalence Levels of Literary Corpus Translation Using a Freeware Analysis Toolkit*