

The Comprehensibility of Readable English Texts and Their Back-Translations

Yan Mujiyanto¹

¹ English Department, Faculty of Languages and Arts, Semarang State University, Indonesia

Correspondence: Yan Mujiyanto, English Department, Faculty of Languages and Arts, Semarang State University, Kampus Sekaran Gunungpati Semarang 50229, Indonesia. E-mail: yanmujiyanto@gmail.com

Received: January 22, 2016 Accepted: February 10, 2016 Online Published: March 28, 2016

doi:10.5539/ijel.v6n2p21 URL: <http://dx.doi.org/10.5539/ijel.v6n2p21>

Abstract

This paper presents the results of a study initiated by the potential employment of readability measures to assess the equivalence of reading ease and grade level indices between source texts and their translations as well as back-renderings. It was questionable whether there was a causal relation between the indices and their comprehensibility levels, because whereas the former concentrated merely on quantities of linguistic elements and their formal relations, the latter considered such factors as particular characteristics of each element, meaning coverage, and readers' socio-psychological background. This study aimed to disclose the relation between the readability measures and the comprehensibility levels of source texts and their translations, as well as back-renderings. A number of English texts, along with their translations in Indonesian, were deliberately chosen for that purpose. The translations were then back-rendered to the source language utilizing *Google Translate*. Comparison between the source texts and their translations as well as back-renderings was capable of showing their similarities in the readability levels and average number of characters, words, sentences, and words per sentence in the texts. And asking prospective readers about their perception concerning their understanding of such texts was capable of disclosing the causal relation between the readability and the comprehensibility levels of the texts.

Keywords: back-translation, comprehensibility, grade-level, readability, reading-ease

1. Introduction

The concept of readability has been widely discussed since it was firstly introduced decades ago. It is as if this concept was the only crucial issue to consider as the starting point of what reasons a text was deliberately written for. A lot of research on this issue has been conducted, viewing it from such perspectives as setting, purposes, and psychological as well as sociocultural aspects. According to DuBay (2004, p. 2), by the 1980s there had been approximately 200 readability formulas and over a thousand studies on the implementation of the formulas attesting to their theoretical and statistical validity. Among the bulk of formulas, the Flesch Reading-Ease test, firstly introduced in 1950s, is still regarded as the one which has been extensively attended. The particular test claims that the purpose of readability measures is to disclose reading ease levels of certain texts; higher scores indicate higher reading ease whereas lower numbers mark that the passages are more difficult to look through.

Besides the test employed for the above purpose, a number of instruments have been applied to determine the grade-levels of certain texts. Among such instruments is the Flesch-Kincaid Grade-level test which—along with four other grade-level tests: the Gunning-Fog Score, the Coleman-Liau Index, the Simple Measure of Gobbledygook (SMOG) Index, and the Automated Readability Index (ARI)—has been implemented to measure to which grade-level of readers a text is normally appropriate. This particular test has been constructed on the basis of the average number of syllables per word and the number of words per sentence. The Gunning-Fog index, developed by Robert Gunning in 1952, estimates the years of formal education required to comprehend a text on a first reading. Meanwhile, the SMOG Index—developed by G. Harry McLaughlin in 1969—has been used for the similar purpose, but this formula has been considered being more accurate than the other ones. The ARI has been made to produce an approximate representation of the United States grade-level needed to comprehend a text. This index mainly relies on the factor of characters per word, considering that it is often faster to calculate. The Coleman-Liau index—designed by M. Coleman & T. L. Liau in 1975—by the same token, relies on characters given that they are more readily and accurately counted by computer programs than

are syllables (Maksymyski, Gutermuth, & Hansen-Schirra, 2015; Readability Formulas; DuBay, 2004; Mujiyanto, 2015).

A number of studies have been conducted in order to uncover the benefits of such measures and indices in the field of language education, particularly in teaching intensive reading. For example, Evanciew & Jones (1996) evaluated several textbooks used in secondary and higher technology education programs relating to readability scores (grade-level equivalences), human interest, and writing style. Meanwhile, Kolahi (2012, pp. 238-239) used the Gunning-Fog Index to measure the readability level in order to demonstrate that Persian translations of English textbooks were less readable than their English counterparts. Wolfer (2015, pp. 36-37) outlined that the main goal of readability studies was “to devise formulae that can be used to directly measure the readability of a text using text surface properties such as mean length of words or sentences.”

Enlightened by years of empirical evidence on the “pitfalls of the readability formulas” (Klare, 1954), a number of experts were triggered to reconsider the reliability of such formulas as means of measuring people’s understanding of certain texts. Wolfer (2015, pp. 36-37), for instance, claimed that the level of observation for a readability formula was always one text as a whole; one specific text only had one index value. Meanwhile, Stephens (2000) stated that readability tests could only measure the surface characteristics of a text; qualitative factors like difficulties in understanding vocabulary, composition, sentence patterns, concreteness, abstractness, obscurity, as well as incoherence could not be measured mathematically. Furthermore, Stephens (2000) viewed that a reading material which received a low grade-level score might be incomprehensible. In other words, while the readability score of a text could be low, its comprehension would be lacking. DuBay (2004, p. 42) pointed out that even though readability formulas were easy to use and capable of indicating the presence of lengthy sentences, they could not be implemented to measure comprehensibility; sentences of the same length might vary greatly in actual comprehensibility.

In order to ascertain the readers’ involvement in understanding verbal texts, be they highly readable or not, experts have been striving to disclose the presence of comprehensibility levels besides that of readability-ease. Quoting Harrison (1980) and Jones (1997), Kolahi (2012, p. 347) showed a significant difference between readability and comprehensibility. While the former is an attribute of text and focuses on textual difficulty, the latter is an attribute of reader concerned about the interaction among such variables as text, task, reader, and strategy. He added that readability might result in comprehensibility in that it was a prerequisite for comprehension. Furthermore, Doherty (2012, p. 93) viewed that readability (defined in terms of linguistic elements) was operationalized as a text-dependent attribute, whereas comprehensibility (characterized as the extent to which a text was understandable) was classified as an attribute of the text which was reader-dependent. The relation between text readability and its comprehensibility, along with the respective text and reader, is illustrated in Figure 1.

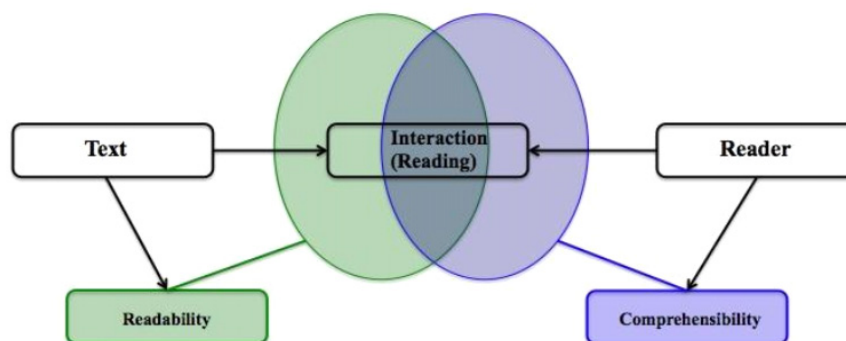


Figure 1. Interaction of attributes of text and reader (Doherty, 2012, p. 93)

Facing such dichotomous perspectives, a number of researchers shift their attention to finding out variables advocating comprehensibility. For instance, Wolfer (2015, pp. 36-37) was interested in studying the comprehensibility of chunks of texts to answer the basic question: “Which linguistic constructions are too complex for which audience?” In order to answer such a question, quoting Schriver (1989) who classified methods of text comprehensibility assessment into three basic categories: text-focused method, expert-judgment-focused method, and reader-focused method, Göpferich (2009, p. 32) preferred the last method

which undoubtedly provided the most reliable results on text comprehensibility because it was dependent more on the audience, whose comprehension problems were the focus of its evaluation. Implementing such a method, Göpferich (2009, p. 40) found that text comprehensibility could be considered from the perspectives of “the cognitive sciences, educational psychology, linguistics, communication theory, and semiotics.” He suggested that comprehensibility analysis could be focused on four comprehensibility dimensions of ‘structure’, ‘concision,’ ‘motivation,’ and ‘simplicity’ as well as two further dimensions of ‘correctness’ and ‘perceptibility.’ He added that the text comprehensibility was determined by the correctness of a text and the ease or difficulty with which it could be perceived and transferred to the reader’s cognitive systems. In addition, Calixto (2016) emphasized that comprehension was ordinarily dependent on a wide range of “perceptual, linguistic, and cognitive processes”.

A number of researchers have employed the concept of readability measures and comprehensibility perspectives in the field of (back)-translation. Hovy (2003, p. 59), for example, found that readability focused on naturalness whereas comprehensibility led to ease of understanding. As an aspect of translation, readability, which has been the subject of numerous machine translation evaluations so far, has also been called fluency, clarity, intelligibility, quality, etc. while comprehensibility has focused on “the extent to which valid inferences can be drawn by combining information from different parts of the document”. Van Slype (1979, p. 62) maintained that even though the comprehensibility of a translation was subject to its readers, “it related to the degree of perfection with which a complete translation could be understood.” In the meantime, Hansen (2015, p. 60) believed that the only reason to translate a text was to make it more comprehensible for the intended recipients. Furthermore, quoting Hönig (1998), Jensen (2015, p. 164) suggested that comprehensibility was an integral part of translation quality assessment. However, amidst the claim that comprehensibility was the central focus of translation research and practices especially in relation to translation quality assessment, the concept had received limited research attention.

To mention a few studies on the importance of comprehensibility in translation, Miller (2001), who contrasted comprehensibility vs. fidelity, showed that even at a lower fidelity rate translation could be more comprehensible although a text with very low fidelity would lead to undesirable task performance. House (2006) reported on a project applying the model of translation evaluation which was designed to answer the question of whether the English language changed textual norms in other European languages and the well-known lexical import from English into other languages.

In a survey on the speedy comprehensibility testing of the worldwide harmonized system of classification and chemicals labeling, UPERDFI (2006) discovered that people regarded labels as the primary source of information despite the fact that symbols played an important role in hazard communication and were keys to attracting attention and the comprehensibility of symbols was highly variable. Meanwhile, Doherty et al. (2010) discovered that the gaze time and fixation count had correlated well with the evaluators’ judgments for the segments used in the study and that the use of eye tracking for automatically evaluating the readability and comprehensibility of machine translation data was worthy of further investigation.

Akamatsu (2011) proved that Trust Rank algorithm was an effective method to detect easy web pages for comprehensibility measure of web documents, whereas Maney et al. (2012, p. 6) found how certain linguistic permutations, omissions, and insertions affected the understanding of translation texts and how deletion of adjectives or verbs caused a significant decline in comprehensibility. Crosbie (2013), investigating the potential of automatic translation engines to be used as a tool for literary analysis, found that many stylistic features were retained in the translation although more subtle features of the texts were lost in translation processes, suggesting that this process might be useful as a preliminary technique in profound studies. Assuming that “broken cohesive chains affected the comprehensibility of translation texts,” Askarieh (2014) found that the influence of broken chains on the comprehensibility of translation was closely related to the effect of the common errors on the translation. As a result, the errors caused cohesive chains to be broken.

The unavailability of study relating readability measures to comprehensibility levels of source texts and their translations as well as back-renderings has encouraged this research to question whether the readability-ease and grade-levels of English texts and their back-renderings have something to do with their comprehensibility levels. Based on the question, this research aimed to uncover (1) whether reading ease levels of source texts are comparable to their respective back-renderings, (2) whether the average grade-levels of the two sets of texts are also comparable, (3) whether character, word, sentence, and words per sentence counts of the text pairs match each other, and (4) whether reading ease of English texts and their back-renderings match their levels of comprehensibility.

2. Methodology

Five English books, along with their translations in Indonesian, were deliberately drawn as samples for this study. They were comprised of two classic novels, i.e., Steinbeck's *The Grapes of Wrath* (1939; Trans. Damono, 2000) and Bronte's *Wuthering Heights* (1847; Trans. Haryo, 2011); two popular novels, i.e. Rowling's *Harry Potter and the Deathly Hallow* (2007; Trans. Srisanti, 2008) and Brown's *Deception Point* (2001; Koesalamwardi & Tanajat, 2006); and an academic textbook, i.e., Armstrong's *A History of God* (1993; Trans. Amm, 2001).

Extracts consisting of approximately 3,000 words each were drawn from the samples, resulting in five pairs of English source texts (ST) and their Indonesian target texts (TT). The Indonesian texts were then back-rendered to English implementing *Google Translate* (<https://translate.google.com>>translate). It yielded five sets of back-translation texts (BT). The results were then slightly edited for punctuation, spelling, grammar, and untranslatability before being put side by side with their respective source texts. Amidst its shortcomings and relatively poor reliability, this machine had been extensively used in Indonesia as a means of word-for-word translation processes in the effort of intensive translation practices which were mostly done manually. The implementation of this program for back-renderings in this study produced back-translation texts which were assumed to be representative of the target texts so that comparing the back-translation results to the source texts would likely provide information about the nature of both the source texts and the target ones.

The readability-ease of the texts was measured by utilizing the Flesch Reading Ease test, whereas their grade-levels were measured using the five instruments mentioned in the introductory section. The application of each of the five instruments produced average grade-levels indicating to what academic levels of readers the texts were supposed to be mostly appropriate. It is assumed that the results of such tests negatively correlate to the Flesch Reading Ease index; a text with a comparatively high score on the Reading Ease test should have a lower score on the Grade-level test.

A questionnaire was made to reveal the respondents' perceptions about the comprehensibility level of the texts. The questionnaire consisted of chunks of the source texts, the back-translations, and statements about the respondents' perceptions of the texts' comprehensibility levels. The variables of comprehensibility levels include (1) comprehension aspects including general ideas, detailed ideas, purposes, intentions, and writer's attitudes contained in the selected texts, and (2) formal features covering grammar use, vocabulary, word length, sentence length, and inter-clause relations (for the complete Questionnaire, see the Appendix). Each chunk was to be judged by perceiving whether it was very easy, easy, moderate, difficult, or very difficult to comprehend. The questionnaire was then exposed to the respondents, i.e., 50 prospective readers, i.e. students of the English Department of Semarang State University, Indonesia. These students were in their 6th semester, attending such courses as Stylistics, Indonesian-English Translation, Academic Writing, and Research Methods in Linguistics. They were in fact appropriate readers for whom the texts in this study were supposed to be produced.

Relating the Flesch Reading Ease measure, the average grade-level indices, and the comprehensibility levels of the texts, this study aimed to compare: (1) the reading ease levels between the source texts and their respective back-translations, (2) the average grade-levels between the two sets of texts, (3) the character, word, sentence, and words per sentence counts of the text pairs, and (4) the causal relation between the reading ease index and the comprehensibility level of the English texts and their back-renderings.

3. Results and Discussion

The quantitative results of implementing Flesch Reading Ease test, the average of the five different grade-level indices, and the comprehensibility questionnaire of the five pairs of texts were presented in Table 1. Employing the first test, it was found that the reading ease level of each of the five text pairs extracted from the five different sources were generally similar to each other, with a tendency that the scores of the back-translation texts were a bit lower than those of the source texts. Then, implementing the grade-level measures, it was found that the grade levels of the back-translation texts tend to be lower than those of the source texts, even though it was factual that the grade level in *Deception Point* (DP) was reckoned a little higher than the rest by four of the five measures causing the average grade level of the back-translation texts in this case to be significantly higher than that of the source texts. Table 1 also shows that the comprehensibility levels of the back-translation texts tend to be significantly higher than that of the source ones, except that in *Wuthering Heights* (WH). In other words, while the reading ease and the grade levels of the back-translation texts tend to be lower than that of the source texts, the comprehensibility levels of the back-translation texts were higher than those of the source texts. It implies that while the relation between the reading ease and the grade levels of the texts are noteworthy, the relation between the two measures and the comprehensibility level on the other is not significant. The comparison of the reading ease and average grade levels between the source texts and their back-translations are

illustrated in Figure 2 and Figure 3 respectively.

Table 1. Reading ease, grade-level, language-element counts, and comprehensibility indices

| Reading Ease, Average Grade Indices, and Comprehensibility Level | GW | | WH | | DH | | DP | | HG | |
|--|------|------|------|------|------|------|------|------|------|------|
| | ST | BT | ST | BT | ST | BT | ST | BT | ST | BT |
| Flesch Reading Ease | 92.8 | 92.3 | 59.7 | 73.8 | 73.8 | 73.2 | 72.5 | 71.8 | 48.4 | 49.6 |
| Flesch-Kincaid Grade-level | 3.1 | 2.9 | 10.2 | 6.7 | 6.5 | 6.4 | 5.5 | 5.8 | 12.5 | 12.1 |
| Gunning-Fog Score | 5.3 | 5.3 | 13 | 9.4 | 8.3 | 8.1 | 7.9 | 8.2 | 15 | 14.3 |
| Coleman-Liau Index | 6 | 6.4 | 10.9 | 9.6 | 10.4 | 10.7 | 10.8 | 10.8 | 12 | 11.6 |
| SMOG Index | 3.7 | 3.9 | 9.1 | 6.8 | 6.7 | 6.8 | 6.2 | 6.5 | 11 | 10.5 |
| Automated Readability Index | 1.6 | 1.4 | 10.6 | 6.4 | 6.7 | 6.6 | 4.6 | 5 | 12.8 | 12.1 |
| Average Grade Index | 3.9 | 4 | 10.8 | 7.8 | 7.7 | 7.7 | 7 | 7.3 | 12.7 | 12.1 |
| Comprehensibility Level | 61.2 | 68.7 | 65.1 | 63.9 | 62.4 | 67.2 | 65.2 | 68.6 | 60.0 | 86.4 |

Note: GW=*The Grapes of Wrath*; WH=*Wuthering Heights*; DH=*Deathly Hallow*; DP=*Deception Point*; HG=*A History of God*; ST=Source Texts; BT=Back-translation Texts.

Figure 2 shows that in general the reading ease level of the source texts is similar to that of the back-translations; it is only the second text, i.e. WH, whose reading ease level of the back-translations is well higher than that of the source texts. It implies that the back-translation texts are much easier to read than the source ones. Figure 2 also shows that the reading ease level of *The History of God* (HG), which is an academic textbook, is the lowest among the five pairs of extracts. It implies that fictional texts are commonly less difficult to read than academic books.

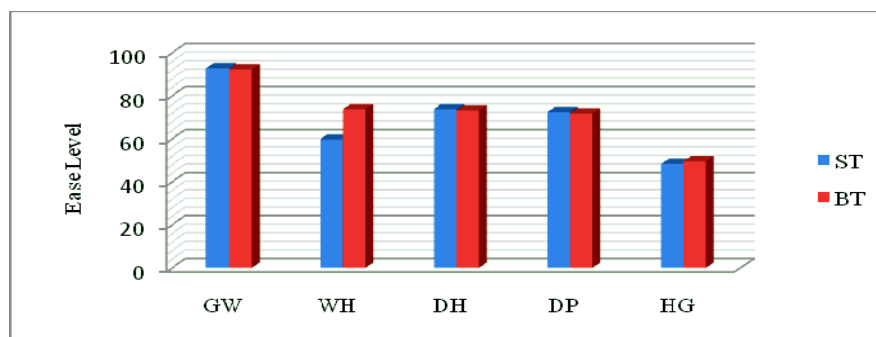


Figure 2. Comparison of the reading ease level between the source texts and their back-translations

The similarity of the reading ease levels between the source texts and their back-translations entails that, to some extent, the translation has achieved certain equivalence of reading ease levels: a source text with a higher level of reading ease is equivalent to its back-translation counterpart or the other way round. If the back-translation text is formally equivalent to the respective translation one, it can be analogically inferred that the source texts and the target texts are commonly equivalent when viewed from the readability perspective. The fact that the reading ease level of WH's back-translation is distinctly higher than that of the source texts implies that literary translation may be focused more on meaning conveyance rather than on the achievement of formal equivalence. The fact that the reading ease of HG is the lowest among the five different texts implies that in general academic texts take more efforts to read than fictional ones in that it takes prospective readers with higher academic level to be able to read such a text appropriately.

Considering the readers' average grade-levels of the five different texts, Figure 3 illustrates that *The Grapes of Wrath* (GW) scored the lowest, whereas HG scored the highest. The figure shows that it was only the source text of WH whose average grade-level necessitated much higher grade-level than the respective back-translations. It also shows that the non-fictional text, HG, scored the highest among the five texts, implying that it takes a higher grade-level of readers to be capable of understanding such a text without significant difficulties.

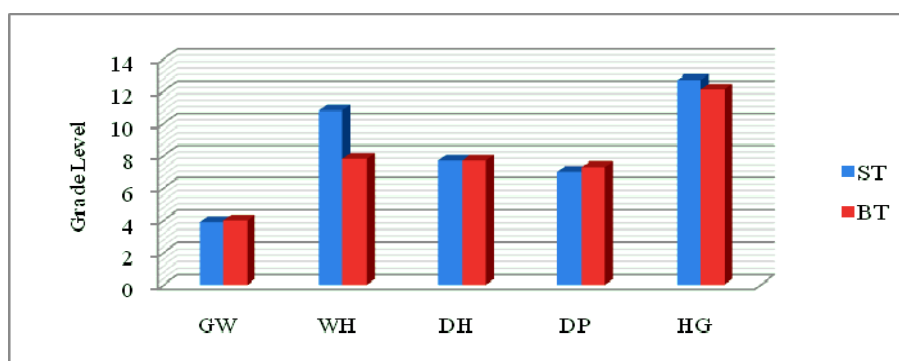


Figure 3. Comparison of average grade-levels between the source texts and their back-translations

Comparing the character, the word, the sentence, and the average number of words per sentence counts between the five pairs of texts, it was found that by and large the number of such linguistic elements in the back-translation texts exceeded those in the source ones. As shown in Table 2, it is only in *The Deathly Hallow* (DH) did the character as well as the word counts of the source texts exceed those of the translations. In DP the number of words per sentence count of the back-translation texts went above those of the source ones. It implies that target texts tend to be longer than their respective source ones. The difference may be dependent on text types and language styles used by the translators when rendering the English texts into Indonesian, taking for granted that the back-translation did not significantly change the number of linguistic elements except the number of characters per word. There is a tendency that words in the Indonesian texts consist of more characters than their counterparts in English.

Table 2. Comparison of character, word, sentence, and words per sentence between source texts and their back-renderings

| Character, word, sentence, and word per sentence counts | GW | | WH | | DH | | DP | | HG | |
|---|-------|--------|-------|-------|--------|-------|-------|--------|--------|--------|
| | ST | BT | ST | BT | ST | BT | ST | BT | ST | BT |
| Character Count | 9,225 | 13,983 | 8,897 | 8,907 | 10,159 | 9,977 | 9,304 | 10,479 | 11,909 | 14,035 |
| Word Count | 2,487 | 3,713 | 1,964 | 2,066 | 2,278 | 2,216 | 2,059 | 2,322 | 2,523 | 3,011 |
| Sentence Count | 224 | 365 | 92 | 137 | 160 | 163 | 217 | 225 | 105 | 130 |
| Words-per-Sentence | 11.1 | 10.2 | 21.3 | 15.1 | 14.2 | 13.6 | 9.5 | 10.3 | 24 | 23.2 |

The comparison of the average number of words per sentence counts between the source texts, the target texts, and the back-translations is shown in Table 3. As shown in the table, it is only in DP did the average number of words per sentence in the back-translation texts exceed those of the source ones. This may be caused by the use of such function words as determiners and auxiliaries in English which rarely have any counterpart in Indonesian. The similarity of the number of words per sentence between the target texts and their back-renderings means that *Google Translate* tends to use formal word-for-word strategy rather than pragmatic or metafunctional one in rendering the translation texts back to English.

Table 3. Average number of words per sentence

| Number of words per sentence | GW | WH | DH | DP | HG |
|------------------------------|------|------|------|------|------|
| ST | 11.1 | 21.3 | 14.2 | 9.5 | 24 |
| TT | 9.4 | 14.9 | 12.5 | 9.4 | 21.7 |
| BT | 10.2 | 15.1 | 13.6 | 10.3 | 23.2 |

Note. ST=Source Texts; TT=Target texts; BT=Back-translation Texts.

The result of comparing the word count contained in the five source texts and their respective back-renderings is shown in Figure 4. In general, the number of words contained in the back-translation texts was higher than that of the source texts. In the case of GW, the number of words contained in the back-renderings was much higher

than that of the source and the target texts, causing the translation as a whole to be much thicker than the source ones. Besides that, even though the number of words in the target texts was generally lower than that in both the source texts and the back-translations, the one in the target texts of the HG was significantly higher than that in both the source texts and the back-translations. It can be implied that although the choice of words may be determined by who the writer or translator is, it may also be influenced by the type of texts to be translated. GW was written and translated by men of letters, whereas HG was prepared by writers who can be categorized as laymen in the field of literature.

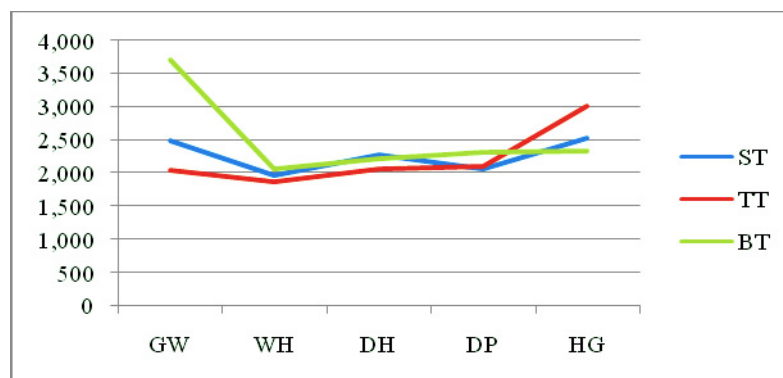


Figure 4. Comparison of word count among the source texts, the target texts, and their back-translations

Similar to the finding on word count, the finding on sentence count also shows that the number of sentences in the back-translations tends to be higher than the one in the source texts. Figure 5 shows that although the sentence counts of the source, the target, and the back-translation of the five different texts were in most cases similar, the sentence count of the GW's back-rendering was significantly higher than the respective source and target texts, signifying that translation yielded by men of letters tends to be larger than the source ones.

The differences in the word and sentence counts between the source texts and their translations indicate the different implementation of translation techniques. On the one hand, the translator implemented formal or word-for-word as well as sentence-for-sentence translation technique, yielding similar number of words and sentences between the source texts and their translations. On the other hand, the translator may apply an idiomatic translation technique tending to produce different numbers of words and sentences because this technique is focused more on conveyance of meaning and intention rather than on form.

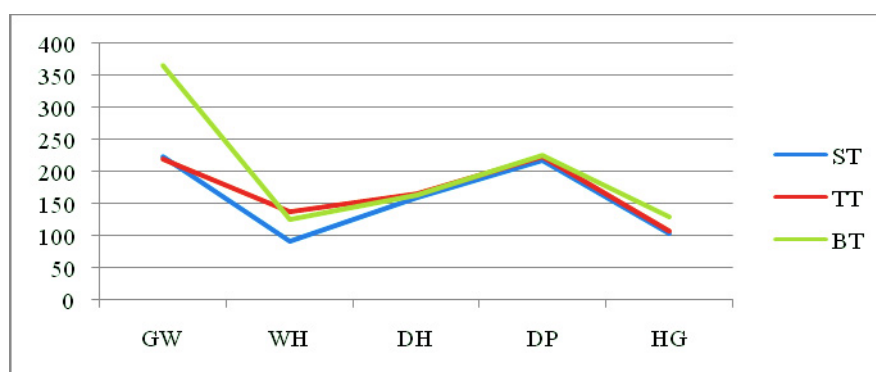


Figure 5. Comparison of sentence count among the source texts, the translation texts, and their back-translations

Comparing the average number of words per sentence contained in the five texts, along with their back-renderings, it was found that the average number of words constructing each sentence in the source texts was similar to that in their back-translations. Figure 6 shows that in the case of WH, sentences in the source texts were averagely constructed of more words than their counterparts in the back-renderings, implying that there is a tendency for the sentences in the back-translation texts to be shorter than those in the source ones.



Figure 6. Comparison of words per sentence counts among the source texts, the target texts, and their back-translations

The comparison of the reading ease indices and the comprehensibility levels between the source texts and the back-translations is illustrated in Figure 7. It can be seen in the figure that the reading ease index of the source texts is higher than its comprehensibility level except in DP, whose comprehensibility level is well above the readability index. Similarly, the comprehensibility level of the source texts is also higher than its comprehensibility level except in the DP, whose comprehensibility level is much higher than its respective back-translations. In other words, the reading ease index and the comprehensibility level of the back-translations tend to be lower than those of the source texts. Figure 7 also shows that the reading ease index and the comprehensibility level of the back-translations are normally higher than those of the source texts, implying that target texts, as represented by the back-renderings, are commonly less difficult to understand than the source ones.

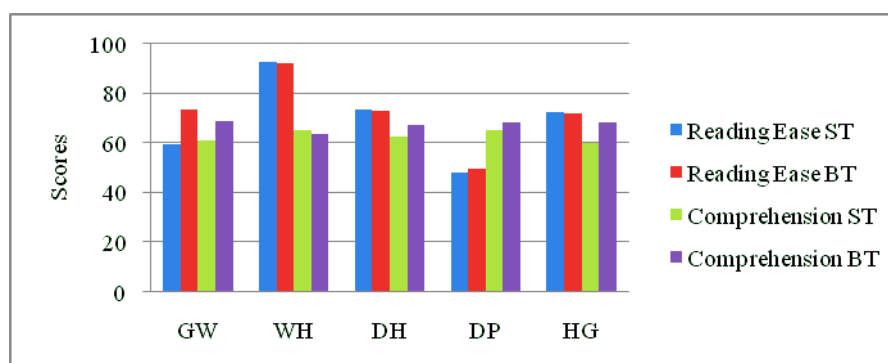


Figure 7. Comparison of reading ease and comprehensibility levels between source texts and their back-translations

Implementing the Two-factor ANOVA with Replication program available at *Microsoft Excel*, it was found that for *Sample*, (i.e., the relation between the readability ease and the comprehensibility level of the source texts and the back-translations), the F-value was as high as 9.046, while the F-critical based on df (0.05, 1,10) was merely 4.965, the p-value of which was 0.013. Because the F value (9.046) is more than the F-critical (4.965) or the p-value (0.013) is less than alpha (0.05), it can be inferred that the readability ease and the comprehensibility levels of the source texts is significantly different from those of their back-renderings. In other words, the different attribution of readability and comprehensibility, as it was claimed by Doherty (2012, p. 93), has caused each of the two measures to determine its own results. A text whose readability ease is low may be easy or difficult to understand, depending on the comprehension aspects and formal features contained in it.

Table 4. Results of Two-Factor ANOVA with Replication

| Source of Variation | SS | df | MS | F | P-value | F-critical |
|---------------------|----------|----|---------|--------|---------|------------|
| Sample | 164.394 | 1 | 164.394 | 9.046 | 0.013 | 4.965 |
| Columns | 870.747 | 4 | 217.687 | 11.978 | 0.001 | 3.478 |
| Interaction | 1085.960 | 4 | 271.490 | 14.939 | 0.000 | 3.478 |
| Within | 181.734 | 10 | 18.173 | | | |
| Total | 2302.835 | 19 | | | | |

Based on the Analysis of Variance (ANOVA) output for *Columns* (i.e., the five different texts), the F-value was 11.978, while the F-critical based on df (0.05, 4, 10) was only 3.478 with p-value of as much as 0.001. Because the F value was higher than the F-critical or the p-value was lower than the alpha, it can be inferred that there is a significant difference among the average scores of the source texts and their back-translations. It implies that the ease or difficulty in comprehending a text, no matter whether its readability level is high or low, is partly determined by the text types. As seen in Figure 7, the reading ease and comprehensibility levels of the source texts and their back-translations vary from one text type to the others. It seems that the possible similarity in reading ease and comprehensibility level among the texts appears at random.

Based on the ANOVA output of the readability ease and the comprehensibility levels of the source and back-renderings of the five different texts, the F-value was 14.939, much higher than the F-critical (3.478) based on df (0.05, 4, 10), or the p-value (0.000) was lower than the alpha (0.05). Therefore, there were interactions between the text types, the readability ease, and the comprehensibility levels of the source texts and their back-renderings. It implies that different text types tend to determine levels of readability ease as well as comprehensibility level, whether it deals with source texts or their back-renderings.

4. Conclusion

The translation of English texts into Indonesian generally results in readability-ease and grade-level equivalence. The difference between the two measures is due to variance in text types. Academic texts tend to be lower in readability-ease and thus higher in grade-level compared to fictional texts.

The fact that the number of linguistic elements including character, word, sentence, and average number of words per sentence in the back-translation texts exceed those in the source texts implies that translation produces longer texts than their sources, no matter whether they are fictional or academic. The disparity in character and word counts of the source texts and their back-translations between *the Deathly Hollows* and the rest of the samples might be caused by the writers' and translators' different use of translation techniques, styles, or rhetoric.

The significant difference in reading ease and comprehensibility levels between the source texts and their back-renderings implies that a text with higher reading ease or lower grade level does not automatically cause it to be easily comprehensible. While reading ease and grade levels are mainly determined by linguistic element counts, comprehensibility is mainly determined by the readers' mastery level of comprehension aspects and formal features contained in the texts.

The deliberate use of back-renderings to represent target texts in this study is supposed to make the source texts more comparable to its translation, assuming that *Google Translate* relies more on formal word-for-word rendering rather than pragmatic or metafunctional bases.

This research was meant to be a preliminary study resulting in a relatively fervent assumption that text comprehensibility levels are truly attributed to their readers whereas readability is more text-dependent. Therefore, it is suggestible to replicate this study exploring readers' real understanding, rather than their perceptions, about the translation of various texts, along with their back-renderings in different settings and comprehension variables.

Acknowledgments

The author would like to thank the Management of the Faculty of Languages and Arts of Semarang State University Indonesia and the students of the English Department at the Faculty who had made this research possible.

References

- Akamatsu, K., Pattanasri, N., Jatowt, A., & Tanaka, K. (2011). Measuring Comprehensibility of Web Pages Based on Link Analysis. *Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE/WIC/ACM*

- International Conference on*, 1. <http://dx.doi.org/10.1109/WI-IAT.2011.242>
- Armstrong, K. (1993/2001). *The History of God* (Z. Am, Trans.). Bandung: Penerbit Mizan.
- Askarieh, S. (2014). *Cohesion and Comprehensibility in Swedish-Translated Texts*. Master's Thesis Department of Culture and Communication Linköping University.
- Bronte, E. (1847/2011). *Wuthering Heights* (Haryo, A. R.B., Trans). Bandung: Qanita.
- Brown, D. (2001/2006). *Deception Point* (Koesalamwardi, I. B., & Tanaja, H. M. Trans.). Jakarta: Serambi Ilmu Semesta.
- Calixto, B. J. (2016). Some Aspects of EFL Reading Comprehension. *Expectative*. Retrieved from E-Revista.Unioeste. Br/Index.Php/Expectativa/Article/Download/530/441
- Crosbie, T., French, T., & Conrad, M. (2013). Stylistic Analysis Using Machine Translation as a Tool. *International Journal for Infonomics (IJI)*, 1(1). Retrieved from <http://uobrep.openrepository.com/uobrep/handle/10547/333145>
- Doherty, S. (2012). *Investigating the Effects of Controlled Language on the Reading and Comprehension of Machine Translated Texts: A Mixed-Methods Approach*. Thesis submitted for the degree of Doctor of Philosophy School of Applied Language and Intercultural Studies Dublin City University.
- Doherty, S., O'Brien, S., & Carl, M. (2010). Eye Tracking as an MT Evaluation Technique. *Machine Translation*, 24, 1-13. <http://dx.doi.org/10.1007/s10590-010-9070-9>
- DuBay, W. H. (2004). *The Principles of Readability*. Costa Mesa, California: Impact Information. Retrieved from <http://www.impact-information.com/impactinfo/readability02.pdf>
- Evanciew, C. E. P., & Jones, K. H. (1996). Using Readability, Human Interest, and Writing Style to Evaluate Technology Education Textbooks. *Tech Trend*, 41(2), 37-38. <http://dx.doi.org/10.1007/BF02818816>
- Göpferich, S. (2009). Comprehensibility Assessment Using the Karlsruhe Comprehensibility Concept. *The Journal of Specialized Translation*, 11, 31-53. Department of Translation Studies, University of Graz. Retrieved from http://www.jostrans.org/issue11/art_goeperich.pdf
- Hansen-Schirra, S., & Gutermuth, S. (2015). Approaching Comprehensibility in Translation Studies. In K. Maksymski, S. Gutermuth, & S. Hansen-Schirra (Eds.), *Translation and Comprehensibility* (pp. 53-76). Berlin: Frank & Timme Verlag für wissenschaftliche Literatur.
- Harrison, C. (1980). *Readability in the Classroom*. New York: Cambridge University Press.
- Hartley, A., Tatsumi, M., Isahara, H., Kageura, K., & Miyata, R. (2012). Readability and Translatability Judgments for 'Controlled Japanese'. *Proceedings of the 16th EAMT Conference*, 28-30 May 2012, Trento, Italy. European Association for Machine Translation. Retrieved from <http://hltshare.fbk.eu/EAMT2012/html/Papers/56.pdf>
- House, J. (2006). Covert Translation, Language Contact, Variation and Change. *SYNAPS*, 19. Retrieved from http://www.nhh.no/Files/Filer/.../House_SYNAPS_19_2006.pdf
- Hovy, E., King, M., & Popescu-Belis, A. (2002). Principles of Context-Based Machine Translation Evaluation. *Machine Translation*, 17, 43-75. <http://dx.doi.org/10.1023/A:1025510524115>
- Jensen, M. N. (2015). Optimizing Comprehensibility in Interlingual Translation: The Need for Intralingual Translation. In K. Maksymski, S. Gutermuth, & S. Hansen-Schirra (Eds.), *Translation and Comprehensibility* (pp. 163-194). Berlin: Frank & Timme Verlag für wissenschaftliche Literatur.
- Jones, K. H. (1997). Analysis of Readability, Interest Level, and Writing Style of Home Economics Textbooks: Implications for Special Need Learners. *Journal of Vocational Home Economics Education*, 12(2), 13-24. Retrieved from <http://www.natefac.org/Pages/v12no2/12-2-13%20Jones.pdf>
- Klare, G. R., & Buck, B. (1954). *Limitations of the Readability Formulas*. New York: Hermitage Books.
- Kolahi, S., & Shirvani, E. (2012). A Comparative Study of the Readability of English Textbooks of Translation and Their Persian Translations. *International Journal of Linguistics*, 4(4), 344-361. <http://dx.doi.org/10.5296/ijl.v4i4.2737>
- Maksymski, K., Gutermuth, S., & Hansen-Schirra, S. (Eds.). (2015). *Translation and Comprehensibility*. Berlin: Frank & Timme Verlag für wissenschaftliche Literatur
- Maney, T., Sibert, L., Perzanowski, D., Gupta, K., & Schmidt-Nielsen, A. (2012). Toward Determining the

- Comprehensibility of Machine Translations. *NAACL-HLT Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)* (pp. 1-7). Montréal, Canada: Association for Computational Linguistics.
- Miller, K. J., Gates, D. M., Underwood, N., & Magdalen, J. (2001). *Evaluation of Machine Translation Output for an Unknown Source Language: Report of an ISLE-Based Investigation*. Washington D.C., USA: The MITRE Corporation.
- Mujiyanto, Y. (2015). *Comparing the Readability Levels of a Source Text and its Back-Translations*. Paper presented at the 4th ELTIL International Conference, Semarang 10-11 October.
- Readability Formulas*. Retrieved from <http://www.readabilityformulas.com/free-readability-formula-tests.php>
- Rowling, J. K. (2007/2008). *Harry Potter and the Deathly Hallows* (Srisanti, L., Trans). Jakarta: Gramedia.
- Schrivver, K. A. (1989). Evaluating Text Quality: The Continuum from Text-Focused to Reader-Focused Methods. *IEEE Transactions on Professional Communication*, 32(4), 238-255. <http://dx.doi.org/10.1109/47.44536>
- Steinbeck, J. (1939/1999). *The Grapes of Wrath* (Damono, S. D., Trans). Jakarta: Yayasan Obor Indonesia.
- Stephens, C. (2000). *All about Readability*. Retrieved from <http://www.plainlanguage.com/newreadability>
- UPERDFI. (2006). Analysis of the Rapid Comprehensibility Testing of the Globally Harmonized System of Classification and Labeling of Chemicals (GHS) Philippines Case Study.U.P. Engineering Research and Development Foundation, Inc.
- Van Slype, G. (1979). Critical Methods for Evaluating the Quality of Machine Translation. Final Report, Bureau Marcel van Dijk/European Commission, Brussels. Retrieved from [http:// www.issco.unige.ch/en/research/projects/isle/van-slype.pdf](http://www.issco.unige.ch/en/research/projects/isle/van-slype.pdf)
- Wolfer, S. (2015). Comprehension and Comprehensibility. In K. Maksymski, S. Gutermuth, & S. Hansen-Schirra (Eds.), *Translation and Comprehensibility* (pp. 33-52). Berlin: Frank & Timme Verlag für wissenschaftliche Literatur.

Appendix

Questionnaire: Readers' Perceptions of the Text Comprehensibility Levels

Instructions:

1. This questionnaire aims to explore the reader's perceptions of the ease or difficulty levels of English texts and their back-translations.
2. Read the following texts carefully in order to understand their meaning.
3. Show your perceptions about the EASE or DIFFICULTY levels by determining that each of the Comprehensibility Elements of the texts is (choose one only):
(1) Very Difficult; (2) Difficult; (3) Moderate; (4) Easy; (5) Very Easy
4. Express your perceptions by providing ticks (V) in the appropriate columns.
5. Thanks for your sincere participation.

| No | Text | Comprehensibility Elements | Reader's Perceptions | | | | |
|----|---|----------------------------|----------------------|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 |
| 1 | I have just returned from a visit to my landlord—the solitary neighbor that I shall be troubled with. This is certainly a beautiful country! In all England, I do not believe that I could have fixed on a situation so completely removed from the stir of society. A perfect misanthropist's heaven: and Mr. Heathcliff and I are such a suitable pair to divide the desolation between us. A capital fellow! He little imagined how my heart warmed towards him when I beheld his black eyes withdraw so suspiciously under their brows, as I rode up, and when his fingers sheltered themselves, with a jealous resolution, still further in his waistcoat, as I announced my name. | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |

| 2 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
|---|---|------------------------|---|---|---|---|---|
| | <p>We immediately got in-home family room that did not have a corridor or living room. They call this the family room as a whole house. Typically, joined the family room with a kitchen, but I suspect that the kitchen in Wuthering Heights is connected with other parts of the building. At least from the sitting room I could hear the clink of soy sauce base of the tongue and eat that intersect.</p> <p>I also do not see signs of burning, boiling, or baking over a large fireplace in the room, or the sheen of copper pans and kitchen utensils of tin that hung on the wall.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 3 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>'Ma,' she said. Ma's eyes lighted up and she drew her attention toward Rose of Sharon. Her eyes went over the tight, tired, plump face, and she smiled. 'Ma,' the girl said, 'when we get there, all you gonna pick fruit an' kinda live in the country, ain't you?'</p> <p>Ma smiled a little satirically. 'We ain't there yet,' she said. 'We don't know what it's like. We got to see.'</p> <p>'Me an' Connie don't want to live in the country no more,' the girl said. 'We got it all planned up what we gonna do.'</p> <p>For a moment a little worry came on Ma's face. 'Ain't you gonna stay with us - with the family?' she asked.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 4 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>Al pressing the gas. "Listen to the engine." Crackling sound-desk is increasingly hard. Tom listened. "Press the gas and then shut down," he ordered. He opened the hood and poked his head into the "Now press the gas." He listened for a moment and then shut down the engine. "Yes, I suppose you're right, Al," he said.</p> <p>"Bearing piston rod, right?"</p> <p>"It seems so," said Tom.</p> <p>"I always love enough oil in there," complained Al.</p> <p>"Yes, just the oil just not up to it. Now it is drier than female monkeys bitch. Well, not-nothing can be done except to remove them. Come on, I run to the front and to the flat place to stop. You run slowly. Do not drop the bowl. "</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 5 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>The two men appeared out of nowhere, a few yards apart in the narrow, moonlit lane. For a second they stood quite still, wands directed at each other's chests, then, recognizing each other, they stowed their wands beneath their cloaks and started walking briskly in the same direction.</p> <p>'News?' asked the taller of the two.</p> <p>'The best,' replied Severus Snape.</p> <p>The lane was bordered on the left by wild, low-growing brambles, on the right by a high, neatly manicured hedge. The men's long cloaks flapped around their ankles as they marched.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 6 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>Both men were sitting in a place that is given to them. Most of the eyes around the table followed Snape and to him that Voldemort spoke first.</p> <p>"So?"</p> <p>"Your honor, the Order of the Phoenix intends to move Harry Potter from a safe place today on Saturday next, at nightfall."</p> <p>Interest around the table sharpened vividly: some people tense, another restless, all staring at Snape and Voldemort.</p> <p>"Saturday ... at nightfall," repeated Voldemort. Red eyes black eyes fixed upon Snape so intense that some people who see them turn away, apparently fearing they would be burned by the ferocity that view.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |

| 7 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
|----|--|------------------------|---|---|---|---|---|
| | <p>In the beginning, human beings created a God who was the First Cause of all things and Ruler of heaven and earth. He was not represented by images and had no temple or priests in his service. He was too exalted for an inadequate human cult.</p> <p>Gradually he faded from the consciousness of his people. He had become so remote that they decided that they did not want him any more. Eventually he was said to have disappeared.</p> <p>That, at least, is one theory, popularized by Father Wilhelm Schmidt in <i>The Origin of the Idea of God</i>, first published in 1912. Schmidt suggested that there had been a primitive monotheism before men and women had started to worship a number of gods.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 8 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>One of the reasons why religion seems irrelevant today is because many of us no longer have a sense that we are surrounded by unseen. Scientific culture has educated us to focus only on the physical world and the material present before us. This method of investigating the world has brought a lot of results. However, one consequence is that we loss of sensitivity of the "spiritual" or "holy" as it covers more traditional community life at every level, and the former is an essential part of the human experience of the world. In the South Sea Islands, they call this mysterious force; others experience it as a presence or spirit;</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 9 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>Toulos Restaurant, adjacent to Capitol Hill, boasts a politically incorrect menu of baby veal and horse carpaccio, making it an ironic hotspot for the quintessential Washingtonian power breakfast. This morning Toulos was busy—a cacophony of clanking silverware, espresso machines, and cell phone conversations.</p> <p>The maitre d' was sneaking a sip of his morning Bloody Mary when the woman entered. He turned with a practiced smile.</p> <p>"Good morning," he said. "May I help you?"</p> <p>The woman was attractive, in her mid-thirties, wearing gray, pleated flannel pants, conservative flats, and an ivory Laura Ashley blouse. Her posture was straight—chin raised ever so slightly—not arrogant, just strong.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |
| 10 | Text | Compr. Elements | 1 | 2 | 3 | 4 | 5 |
| | <p>Rachel sighed in exasperation. Since he had already trying hard not to glance at his watch. "Dad, I really do not have time to call. And I hope Dad would stop trying to—"</p> <p>"You must take the time to do things importantly, Rachel. Without love, everything would be meaningless."</p> <p>Some memories occurred to Rachel, but she chose silence. Looks like, acting like a big no difficult for his father. "Dad, you said you wanted to see me."</p> <p>Dad says there are important things. "</p> <p>"True." Rachel Sexton looked at more closely. Rachel felt her defense partially melted under sharp gaze of his father, so he cursed the man force it is in the heart.</p> | General idea | | | | | |
| | | Detailed idea | | | | | |
| | | Purpose | | | | | |
| | | Intention | | | | | |
| | | Writer's attitude | | | | | |
| | | Grammar | | | | | |
| | | Vocabulary | | | | | |
| | | Word length | | | | | |
| | | Sentence length | | | | | |
| | | Inter-clause Relations | | | | | |

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).