

## A Readability Study and Its Relevance to Simplification on Translations of *Lun Yu*

YU Juan<sup>[a],\*</sup>

<sup>[a]</sup>School of Foreign Languages, Sun Yat-Sen University, Guangzhou, China.

\*Corresponding author.

**Supported by** Guangdong Provincial Philosophy and Social Sciences Project of “12-5” Plan: A Linguistic Study of the Readability of English Original Writings and Their Simplifications (GD13CWW02).

Received 28 September 2014; accepted 5 November 2014  
Published online 26 December 2014

### Abstract

This research will apply readability theory into translation studies of *Lunyu* or “*the Analects of Confucius*” in Chinese and English versions. With the application of readability formulas, the analysis will explain the readability statistics in two ways: the Flesch reading ease score and other Reading Level index. The two translated texts will be compared in terms of four aspects: word numbers, lexical density, sentence numbers and average sentence length so as to find out the different degree of simplification in the translations. The findings show that the readability of Roger Ames’s translated texts is higher than James Legge’s; hence is much more difficult to read. The simplification degree is in accordance with the readability result, which indicates that readability analysis results can be referenced for studying translation features like simplification.

**Key words:** Readability; *Lunyu* English translations; Readability formula; Simplification

Yu, J. (2014). A Readability Study and Its Relevance to Simplification on Translations of Lun Yu. *Studies in Literature and Language*, 9(3), 47-57. Available from: <http://www.cscanada.net/index.php/sll/article/view/5856>  
DOI: <http://dx.doi.org/10.3968/5856>

### INTRODUCTION

Readability is an important concept in the theory of applied linguistics. Zamanian and Heydari (2012) in

*Reability: State of Art*, points out that the definition provided by Dale and Chall (1949) possibly the most comprehensive: “The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it.” (qtd. in Zamanian & Heydari, 2012) There are many factors contribute to readability of a text which including: long- term Proportion; different words: the degree of abstraction; Vocabulary; number of pronouns; number Preposition Number; affixes; the number of difficult words; sentence length. (Betts, 1949; Dechant, & Smith 1961/1977; qtd. in Wang & Yang, 2012)

The above remarks show that readability research in China at present mainly focuses on the teaching of reading and teaching materials compilation evaluation, reading tests and other fields in applied linguistics. At this point, it is necessary to emphasize the application of readability theory, formulas into other areas, like text analysis, and translation studies for two reasons. On the one hand, the target of these areas is text, which can be regarded as the first common ground; on the other hand, application of readability formulas can improve the objectivity and effectiveness of text analyzing, because it is computerized and most of the calculations have been done by softwares rather than human, which avoid the personal intervene of the results.

Drury (1985) summarized two reasons for the population of readability formulas: Firstly, they are objective and can be used as legal and contractual criteria; secondly, the simply counting characteristic of most has made them readily adaptable to computer application.

In other words readability formulas can provide accurate comparison of texts so that assessment can be made between the texts and a reader’s reading abilities and interests.

The popular formulas are: Dale-Chall; SMOG; POG; Fry; Flesch Reading Ease; and so on; among which the Flesch Reading Ease is the most widely used and tested. (Dale & Chall, 1949; Klar, 1963)

The updated Flesch Reading Ease formula is: Score = 206.835 - (1.015 × ASL<sup>1</sup>) - (84.6 × ASW<sup>2</sup>). The interpretation for the scores and the corresponding difficulty and grade levels is illustrated in the following table.

**Table 1**  
**Reading Ease Scores and Estimated Reading Grade of Flesch**

Reading	Style description	Estimated reading	Estimated percent
0 to 30	Very difficult	College graduate	4.5%
30 to 40	Difficult	College grade	33%
50 to 60	Fairly difficult	10 <sup>th</sup> to 12 <sup>th</sup>	54%
60 to 70	Standard	8 <sup>th</sup> to 9 <sup>th</sup>	83%
70 to 80	Fairly easy	7 <sup>th</sup>	88%
80 to 90	Easy	6 <sup>th</sup>	91%
90 to 100	Very easy	5 <sup>th</sup>	93%

Note. Cited in Zamanian and Heydari, (2012).

According to Flesch (1948), the measurement of work length is an indirect way to measure word complexity. And word complexity is an indirect way to measure abstraction. For the same reason sentence complexity and abstraction can be measured in the same way. It can be summarized that Flesch Reading Ease Formula is essentially a measurement of the level of complexity and difficulty of an article.

## 1. LITERATURE REVIEW

### 1.1 Review of Oversea Readability Study

Before we move on the statistics' analysis, it is worthwhile summarizing the long history about the development of readability study. The literature on readability formulas research can be summarized in three types:

- a) History of readability theory and methodology;
- b) Application of readability formulas;
- c) Advantages and disadvantages of utilizing readability formulas.

Researchers of the first type are: Flesch (1948), Dale-Chall (1948) qtd.in DuBay, 2004); Gunning (1952); Fry (1968); McLaughlin, (1969). (qtd. in DuBay, 2004) Second type study focuses on the application of readability formulas, which were originally created to test the readability level of textbooks. Nowadays, the application of readability formula has proven to be an effective tool to test not only textbook, but also other texts, such as prescriptions, mass media texts, website information, legal documents, and so on. Validity

<sup>1</sup> ASL = average sentence length (the number of words divided by the number of sentences).

<sup>2</sup> ASW= average number of syllables per word (the number of syllables divided by the number of words).

and appropriateness are the major issues discussed by researchers like Kirkwood and Wolfe, 1980; Bertram and Newman, 1981; Frase, Rubin, Starr, and Plung, 1981; Bailin & Grafstein, 2001. (qtd. in DuBay, 2004) Such studies give theoretical background and empirical assessment of readability formulas, and can help us to choose the right one when conducting research.

This section reviews and summarizes the major characteristics of the oversea research results and readability research. It can be seen from the review, readability research in Western country has been developed almost a century, and the research system and model have been formed successfully.

### 1.2 Review of Domestic Readability Study

Compared with Western countries, the readability study in China started almost 50 years later and at present, which can be summarized in the following passages.

#### 1.2.1 Textbook Readability Study

Domestic textbook readability researchers focused on two perspectives: First, the use of foreign theoretical evaluation of college English teaching materials readability and suggest improvements; second, the insights of their research on readability analysis. It is common that Chinese researchers are active in introducing foreign readability research findings or discussing the impact of conducting readability analysis on teaching materials as well as the subjects. It is found that Foley readability formula is applied by most domestic scholars readability formulas.

Guo (2010) points out that readability is an important link between the text and the reader. He explores the factors that would influence the readability of Chinese text and developed a computerized formula that can analyze the readability of China for oversea students major in Chinese. In his thesis, Yang (2008) uses traditional questionnaires method to find out the readability level for higher-level Chinese textbooks, and discusses various factors that affect the readability of Chinese textbooks. Xiao (2012) first summarizes the research status of evaluation of textbooks in China, and then carries out an in-depth analysis on text readability as a method of textbook evaluation. She focuses on three text evaluation dimensions: readability formulas evaluation, illustration evaluation table, and material organization chart. The result shows that it is recommendable to understand the pros and cons about each method.

#### 1.2.2 Newspaper Readability Study

Ge (2007) analyzes fanatical newspapers in order to find out whether readability influences the motivation of the investors by using the FLESCH index and variables index of Chinese. The results show that the relation between information and readability is not always in proportion, because information of newspapers in 2005 is increased significantly, but the figure of readability has not increased accordingly.

Based on linguistic features, quotations, background information etc., Jia (2004) analyzes the readability of Chinese news and English news, and explores the factors that affect the readability differences.

### 1.2.3 Readability Research of Vocabulary

Mu (1996) pointed out that the vocabulary is a fundamental factor affecting the readability. He listed eight factors that can affect the readability of reading materials from the perspective of vocabulary, and demonstrates these factors are important reference for compiling and choosing reading materials.

### 1.2.4 Computer Linguistics Readability Study

According to the paper of Zhang (2010), Lin (qtd. in Zhang Ci, 2010) is the first person that studies English readability in China. Xin and Cheng (2010) not only introduces English readability theory, but also proposes a new way to test the key variables that affect the readability of the information. Based on a large number of training data, they build a model<sup>3</sup> to calculate readability based on the measurement information and the use of database technology, which proves to be very accurate in measuring the variable range and other characteristics.

### 1.2.5 Readability and Translation Study

Readability studies in translation area are comparatively rare and there are two types of research. Studies of the first type aim to discuss how to improve translation by applying readability index and readability formulas and the second type often makes a comparative readability analysis between target text and source text. Wang (2012) discusses the relation between difficulty of translation and readability, and claims that readability can be an indicator for translators when they make choices on translation strategies, text styles and grammatical characteristics. Zhou (2012) demonstrates that the possibility of improving translation readability by reducing the English prepositions, articles and pronouns, and proves the feasibility of the strategy and the importance of readability on translation. Li's paper (2006) about the Chinese translation of *Historical Records*, is a work that analyzing the Chinese translation texts by applying readability theory and formulas. It provides a new perspective for researchers to study classical Chinese text and translation features.

In this section, we have reviewed the relevant studies of readability in the literature, including the history, the study focus as well as the achievements. It is interesting to note that, compared with Western readability research, most of the studies conducted by Chinese scholars are confined in research areas like foreign language teaching, second language acquisition, and textbook assessment. In recent years, however, the research range has begun to shift and expand to other areas like literature study, text analysis and translation.

---

<sup>3</sup> ERDA (English Readability & Difficulty Assessment) was an software developed by Xin Fukun & Cheng Donyuan.

## 1.3 Simplification in Translation: An Overview

Mona Baker (1993) advocates a study of the universal features of the translational language by using the comparable corpus. She points out that these features “typically occur in translated texts rather than original utterances” and these features are “thought to be independent of the influence of the specific language pairs involved in the process of translation” (p.243). In other words, the universal features of translation will occur in any translated texts; and they appear naturally in the translation process with or without being controlled by translators. The features identified by Baker (1993) are: explicitation, simplification, and normalization/conservatism. As this study focuses on the discussion of simplification, we will make an overview on this feature in the following passages.

Simplification can be defined as translation strategies and text features produced when translators tend to simplify the translation during the translating process so as to make translation reader-friendly. It is a feature has been observed prior to the development of analytical tools of corpora linguistics. Therefore, we can discuss simplification in two different stages: simplification before the use of the CTS (corpus translation studies) approach and simplification after the use of the CTS approach. Some of the relevant studies will be reviewed as follows.

### 1.3.1 Simplification Studies Before the Development of CTS

Simplification can be described as the tendency, in the translated texts, of using the more simplified language compared with that of in the SL text. Blum-Kulka and Levenston (1983) discuss the meaning of the term “lexical simplification” in the context of second language acquisition. They view lexical simplification as “the process and/or result of making do with less words” (p.119). Blum-Kulka and Levenston (1983) conduct a small-scale empirical study to examine a few aspects of lexical simplification, especially those strategies which play a role in the formation of a learner's interlanguage and its fossilization. Their studies suggested that simplification is a universal feature of language use which may be manifested in a number of linguistic contexts, including the creation of a learner's inter language. The method is used to measure lexical control of non-native learners as compared with native speakers, while the second is intended to measure the learners' ability to recognize lexical items as opposed to their ability to produce them. Results support the hypothesis of lexical simplification in language learners, as well as that of inter language fossilization. Built on the evidence from studies of translation from Hebrew into English and investigations of other types of language mediation involving these languages, Blum-Kulka and Levenston (1983) declare that lexical simplification operates according to five strategies

or principles which derive from the individual's semantic competence in his/her mother tongue.

The first strategy is "the use of the superordinate terms when there are no equivalent hyponyms in the target language" (qtd. in Laviosa, 2002, p.44). The example provided by Blum-Kulka and Levenston is the Hebrew word "almoni", which is translated into an English common word "man", since there is no close lexical equivalent for the word. The second one is "the approximation of a concept expressed in a source language word through a target language expression which does not have the full range of the original meaning" (p.44). This choice is made because of "the lack of culturally matching terms in the two languages" (p.44). Words which contain cultural-specific meanings may not be found in another language and those cultural-specific words might be replaced by general meaning words without specific cultural connotations (p.44). The Hebrew word "hupa", which has strong Hebrew religious and cultural meanings, is translated as "canopy" in English with the general meaning. The third strategy is to replace "infrequent target language words" with reader-friendly, common-level or familiar words to achieve the goal of readability (Laviosa, 2002, p.44). This strategy is used in Biblical translation; thus, "stayed" is replaced with "remained"; "happened" is in place of "changed". The fourth strategy is "transfer of all the functions of a source-language word for its target-language equivalent" (p.44). This strategy is flexible transfer by adding some words or providing some explanations, not getting "too tied to his source language" (Blum-Kulka and Levenston, 1983, p.133). The fifth strategy is "the use of paraphrase to reduce the cultural gap between SL and TL", and "the use of circumlocutions instead of conceptually matching high-level words or expressions, especially with theological, culture-specific or technical terms" (p.44).

Vanderauwera (1985) proposes the use of modern, colloquial, simple and confidential synonyms in place of old, formal, affected and high-level words in the source texts when making the survey of 50 English translations of Dutch novels. Concerning syntactic simplification, Vanderauwera (1985) finds several instances where complex syntax is simplified by replacing non-finite clauses with finite ones and by suppressing suspended periods (Laviosa, 2002, p.47). To demonstrate stylistic simplification, she proposes some strategies such as "breaking up long sequences and sentences, replacing elaborate phraseology with shorter collocations, reducing or omitting repetitions and redundant information, shortening overlong circumlocutions and leaving out modifying phrases and words" (p.47). The aim of these simplification strategies is to make the translation "easier, more coherent, fluent and familiar" (p.47).

Toury (1995) provides an example of the type of transfer noted by Blum-Bulka and Levenston (1983)

when he discusses the word "na'ara" which in Hebrew refers mainly to "a teenager", but the English translation from Hebrew has acquired some functions, such as "girl" (Toury, 1995, pp.209-210). He (Toury, 1995) describes the phenomenon as "a verbal formulation of a translation [that] is partly governed by a felt need to retain aspects of the corresponding source text invariant" (qtd. In Laviosa, 2002, pp.49-50). His study suggests that some words in the source texts tend to be translated into the words with more general meanings.

Klaudy (1996) introduces the concept of "translational operations" (qtd. In Laviosa, 2002, p.48) and defines it as a complex mental operation happening when the mind produces linguistic forms through another language indirectly. Klaudy (1996) observes the phenomenon in multi-language texts such as English, German, French and Russian into Hungarian. In her study, translational operations have been divided into three categories. The first one is "language-specific" operations, which are the process involving different grammatical and lexical structures. The second one is "culture-specific" operations, a process related to various cultural differences in various languages. The final one is "translation-specific" operations, which focus on the process of translation by itself. The above studies are in relation to the simplification feature in translation before the development of the CTS approach. Most of them are descriptive evaluations of translated texts, and the analyses have been carried out manually by using small parallel texts.

In addition, most of the research data are retrieved from the literary texts, and observations mainly focus on the translations from one language to another, not in both directions.

The analysis consists mostly of shifts that occur during the process of translation at the level of the sentence, without regard for directly assessing the influence of simplification strategies on the entire text. Laviosa (2002) points out that their weakness is "lack of a clear definition of universals of translation in general and of simplification in particular, unavailability of a large amount of textual material, and lack of a consistent methodology" (p.51). Therefore, translation studies about the simplification feature with the CTS approach are expected to avoid these defects.

### 1.3.2 Simplification Studies After the CTS

Initiated in 1990s, Mona Baker (1995, 1996) proposed to analyse the existence of universal features of translation, their theoretical plausibility and the feasibility of investigating them systematically by using large scale corpus. Following Laviosa (1998) conducts a study using a corpus of one million words of translated English and a comparable corpus of non-translated English. This study sets out to develop a viable descriptive and target-oriented corpus-based methodology for the systematic study of the nature of the translated text. There were three objectives realized in this study.

a) the elaboration of criteria for designing a monolingual, multi-source language English Comparable Corpus (ECC);

b) The application of these principles to the creation of two sub-sections of ECC, namely newspaper articles and narrative prose;

c) The investigation of simplification as a universal of translation, as a way of testing the viability of the proposed methodology.

The English Comparable Corpus consists of two computerized collections of texts in English: one, referred to as the Translational English Corpus (TEC), includes translations from various source languages; the other, called the Non-translational English Corpus (NON-TEC), comprises original English texts of a similar type that are produced during a similar time span. The investigation of the ECC focuses on global aspects of lexical and stylistic simplification and the result reveals four consistent patterns of lexical simplification in translated versus original texts. These patterns include: a) relatively lower high frequency versus low frequency words; b) relatively greater repetition of the proportion of lexical words versus grammatical words; c) relatively high proportion of most frequent words; d) less variety in the words most frequently used.

Followed this study, Laviosa (1998) uses a comparable corpus to investigate the core patterns of lexical use of English narrative prose. The comparable corpus consists of a subcorpus of 14 translated English works (two biographies and twelve fictions, translated from German, Greek, Romance, Semitic and Slavic languages) and a subcorpus of original narrative texts selected from the British National Corpus (BNC). The findings show that the lexical density is apparently lower in translated narrative while the average sentence length is relatively higher. Laviosa (1998) claims that even excluding the texts with the highest sentence length, the average sentence length for the translated prose is still higher than the comparable original group (18.62 vs. 15.62).

In 2002, Laviosa (2002) conducted another corpus-based study of simplification raising three hypotheses (pp.60-62): a) In a multi-source-language comparable corpus of English, the range of vocabulary used in the translational texts was narrower than the range of vocabulary in the non-translational texts and this difference was independent of the source language variable. b) In a multi-source-language comparable corpus of English, the translational texts had a lower ratio of lexical to run words than the non-translational texts and this difference was independent of the source language variable. c) In a multi-source-language comparable corpus of English, the translational texts had a lower average sentence than the non-translational texts and this difference was not influenced by the source language variable. According to Laviosa (2002), there are three

search criteria to examine the simplification feature: lexical variety, information load and sentence length. She analyzes a comparable corpus consisting of a subcorpus of translated English (newspaper texts and narrative prose translated from several languages) and a subcorpus of non-translated English. Her findings justify the former two hypotheses because the proportion of high frequency words to low frequency words is higher in translated texts than that in non-translated texts; the proportion of content words to grammatical words is lower in translated texts than that in non-translated texts. Laviosa (2002) also finds out that the translator might focus on the creation of some artistic effect or literary value, instead of the simplification of sentences structures in literature translation, because the main purpose of literary texts is aesthetic appreciation; while in writing news articles, the translators tend to adopt shorter sentence so as to help the readers more easily to catch the message and fulfill the purpose of information acquisition.

#### 1.4 Short Review of Studies on *Lun Yu* Translations

There are several systemic functional linguistic studies of *Lun Yu* and its translations as text in the literature to date. For example, Fang (2006) investigates 'the socio-historical and ideological context' for the reproduction of *Lunyu*, with a focus on contextual features, the hierarchical structure and the ways of realizing Confucius' ideas in terms of the lexico-grammatical choices. Wei and Zhang (2006) offers a register analysis and study translated texts by interpreting them within their context of situation. By contrast, Zou's study (2007) focuses on the contexts of culture. And Chen (2009) looks at translated texts from a functional discourse analysis perspective with a focus on the lexico-grammatical stratum. (qtd. In Huang, 2012)

To conclude, the study of *Lun Yu* translation is characterized with multidisciplinary and multi-perspective. However, the study focusing on the readability analysis of *Lunyu* translated texts is rare. This paper, by means of readability formulas and other corpus tools, will analyze the translations in two perspectives: readability level and simplification features. The former is a concept of applied linguistic; the latter is a concept from translation studies. To be more specific, it is a study carried out in two sections: One examines the readability level by applying formulas; the other explores the simplification features in translations and its relevance with the readability level statistics.

This study aims to analyze the translated texts of *Lun Yu* (or *The Analects*) in two perspectives: readability analysis as well as its relevance and implication in the translation simplification feature. In terms of readability study, some formulas are applied so as to find out the basic readability statistics about each text. As for the translation features research, some other corpus tool, such as AntConc, will be used and the analysis of translation

features will mainly base on the readability statistics. It is by means of these that the text can be fully explored.

## 2. READABILITY ANALYSIS OF LUN YU TRANSLATIONS

### 2.1 Research Purpose

As illustrated in the literature review, there are not many studies on the relations between readability analysis and translation. The unpublished thesis written by Fen Qiuxiang (Fen, 2006) can be regarded as an example to discuss the connection of readability study and translation. Based on empirical study, she demonstrated that if the number of English prepositions, articles and pronouns can be reduced, the readability level of the translated text can be increased in her paper and proved there is an important connection between readability level and translation text.

### 2.2 Methodology

The major research tools implemented in the study include:

- a) The Flesch Reading Ease Formula to calculate the sentence length and number of syllables and words, etc..
- b) The Excel tool to calculate the percentage and ratio of the index.
- c) AntConc to check the type-token ratio.

Therefore the research procedures can be listed as follows:

First step: The English text readability analysis was entered into the page to draw relevant readability test data. It should be noted that in the English text cannot be tested on the same page and because there is no website that provides test for Simplified Chinese, so I use the readability test platform developed by National Taiwan Normal University. This website and software are designed to test of readability level of Chinese.

Second step: After getting the data from the website the EXCEL will be applied to calculate the average value of each translation and make the comparison visualized.

Third step: The final stage of this study is the presentation of the statistical results with tables and discussion of research findings. The results of the statistics are tabulated and presented. The charts are made with EXCEL.

Fourth step: presenting the implication of the study and provide the suggestions for the future study.

### 2.3 Findings and Explanation

This part aims to present the statistical findings of the study. It has been introduced that the process of analysis occurs not only in linguistic but also in translation study. Therefore, the findings will involve both of the two aspects. The first part is the general figures of readability report and comparison between the Flesch Reading Scores of the two texts. The second part is an explanation of detailed figures that contributed to the readability result and exploration of the discrepancies of different reading level index.

#### 2.3.1 Statistic Findings of Flesch Reading Ease Score

The Flesch Reading Ease Scale is the most widely used formula outside of educational circles. It is also the easiest formula to use, and it makes adjustments for the higher end of the scale. It measures reading from 100 (easy to read) to 0 (very difficult to read). A zero score indicates text has more than 37 words on the average in each sentence and the average word is more than 2 syllables. Flesch has identified a “65” as the Plain English Score. It also provided an interpretation table to convert the scale to estimate reading grade and estimated school grade.

In our study, we have applied the online resource<sup>4</sup> for automatic readability analysis, and the statistic results are presented in Table 2.

**Table 2**  
**Flesch Reading Ease Score for Each Book Roger Ames’s Translation**

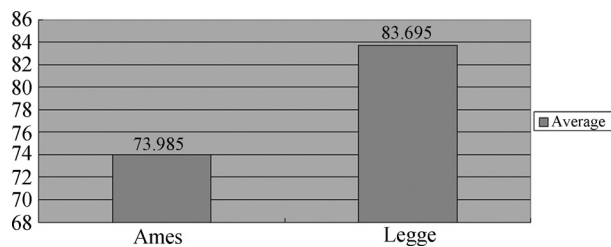
Book	Flesch reading ease score		Style description		Estimated reading grade	
	Ames	Legge	Ames	Legge	Ames	Legge
1	68.9	63	Standard	Standard	8 <sup>th</sup> to 9 <sup>th</sup> grade	8 <sup>th</sup> -9 <sup>th</sup> grade
2	59.1	84.1	Fairly difficult	Easy	10 <sup>th</sup> -12 <sup>th</sup> grade	6 <sup>th</sup> grade
3	75.3	87.3	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
4	73.3	88.1	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
5	80.3	87.5	Easy	Easy	6 <sup>th</sup> grade	6 <sup>th</sup> grade
6	79.4	86.1	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
7	76.6	85.4	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade

To be continued

<sup>4</sup> <http://www.editcentral.com/gwt1/EditCentral.html#notes>

Continued

Book	Flesch reading ease score		Style description		Estimated reading grade	
	Ames	Legge	Ames	Legge	Ames	Legge
8	74.1	81.4	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
9	82	87.6	easy	Easy	6 <sup>th</sup> grade	6 <sup>th</sup> grade
10	77.5	85.5	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
11	81.1	88.1	easy	Easy	6 <sup>th</sup> grade	6 <sup>th</sup> grade
12	66.4	81.5	Standard	Easy	8 <sup>th</sup> -9 <sup>th</sup> grade	6 <sup>th</sup> grade
13	70.5	82.8	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
14	74.9	85.3	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
15	67.9	85.5	Standard	Easy	8 <sup>th</sup> to 9 <sup>th</sup> grade	6 <sup>th</sup> grade
16	70.2	83.2	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
17	75.8	88.1	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
18	82.6	87.1	easy	Easy	6 <sup>th</sup> grade	6 <sup>th</sup> grade
19	71.2	80.3	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade
20	72.2	76	Fairly easy	Fairly easy	7 <sup>th</sup> grade	7 <sup>th</sup> grade
Average	73.985	83.695	Fairly easy	Easy	7 <sup>th</sup> grade	6 <sup>th</sup> grade



**Figure 1**  
**Flesch Reading Ease Score**

From Figure 1 we can see that the average Flesch Reading Ease Scores of Roger Ames’s Translation is 73.985; the reading style is fairly easy. The figures fluctuate from Standard reading to Easy reading. Most of the readability level are between 70-80, and there are two figures indicate standard reading level, Book 1 and Book 12. The most difficult as well as the easiest is Book 12 and Book 18, where the figures are 66.4 and 82.6, respectively.

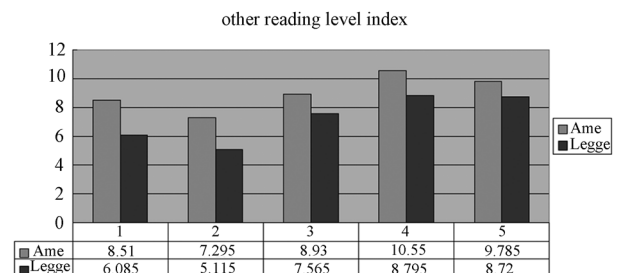
The average Flesch Reading Ease Scores of James Legge’s Translation is 83.695, 10 points higher than Ames’s translation. The corresponding reading style is easy and the figure is comparatively stable. There are only two Chapters that are not at Easy level, the first Book and the last Book. The readability of Book 1 is 63, standard reading level; while the readability of Book 20 is 76, fairly easy style. Therefore, in James’s translation, Book 1 is the most difficult to read and Book 20 is the easiest.

### 2.3.2 Statistics of Reading Level Index

Except the Flesch Reading Ease Scores, we have also present the index of other reading level analysis, such as

the Coleman-Liau index, Gunning fog index, and SMOG index, so as to make the results more convincing.

Figure 2 shows that Roger Ames’s figure of each index is higher than James Legge’s translation, which in accordance with the Flesch score, as it is the rule that the higher reading level index, the less the reading difficulty. The average automated readability index of Ames’s translation is 8.51; Flesch-Kincaid grade level is 7.295, the Coleman-Liao index is 8.93, the Gunning fog index is 10.55, and the SMOG index is 9.785. The average automated readability index of Legge’s translation is 6.085; Flesch-Kincaid grade level is 5.115, the Coleman-Liao index is 7.565, the Gunning fog index is 8.795, and the SMOG index is 8.72.



**Figure 2**  
**Comparison of Other Reading Level Index Between Two Translations**

## 3. DISCUSSION

As far as the readability is concerned, it can be concluded that the translation of James Legge is easier than Roger

Ames's. Nevertheless, challenges to the validity of readability formulas can be troublesome for us to accept the result. In order to make our research more valid, we will explore the relevance between the readability index and other statistics obtained by some other online text analysis tools. They will be used to get the statistical features of a text, like the type-token ratio, lexical and the syntactic information.

### 3.1 Our Assumptions

To be in line with Mona Baker as well as other scholars who share the same ideas, we agree that translation is a distinct linguistic behavior and thus translated texts inevitably have distinct features, which can be observed and measured as they are consistently recurring in the surface structures of translated texts. Based on the statistics that we have obtained in the part of readability study, we have formulized two assumptions in this paper:

A. In terms of the relevance between simplification and readalibility level, we assume that readability can be a reference index to simplification analysis of a text. To be more specific, if we want to know whether a translated text is much simpler than the original text, we can analyze the readalibility level of the texts first.

B. According to the first assumption, we believe that if a translated text is of high readability level, it will present a simplification tendency in lexical as well as syntactic features; while if the readability level of a translated text is low, the lexis and sentences will presented to be more complex than other texts.

In the following passages, we will compare the

statistics of readability features in two layers (lexical and sentence) and examine whether the two assumptions about the relevance between readability level and simplification feature can be justified or not.

### 3.2 Comparison of Simplification Features

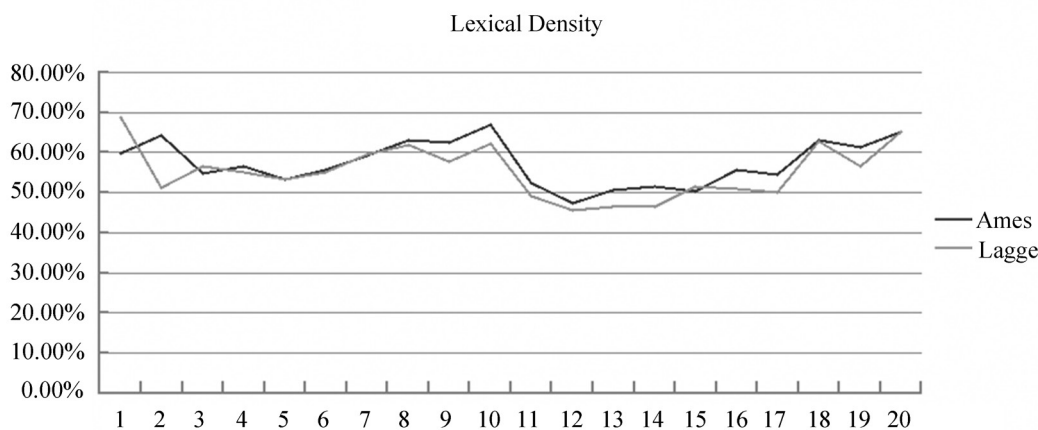
#### 3.2.1 Lexical Simplification

Lexical Density is frequently used by linguistics to analyze simplification in lexical layer, and there ways to get lexical density vary in technology as well working style. For example, websites like [UsingEnglish.Com](http://UsingEnglish.Com) provide online calculation by using the following formula: "LD= (Number of different words / Total number of words)×100".

According to description of the website, any lexically dense text has a lexical density of around 60%-70%, and those around 40%-50% can be regarded as having a lower lexical density.

As a grammarian, M. A. K. Halliday also gives a definition of lexical density, "Lexical density can be measured as the number of lexical items (content words) per ranking clause. (A rank clause is one that is functioning as a clause, independent or dependent; not "rank-shifted" to become part of something else.) Since this is a comparative measure, it does not matter exactly where the line between content words and function words is drawn, provided it is drawn consistently for all the texts under study. (Halliday, 2008, p.158)

As most of the software can calculate the lexical density automatically, we choose to use <http://textalyser.net> as the website to analyze our text and the statistics are presented in Figure 3.



**Figure 3**  
**Lexical Density of Two English Translations**

The line in Figure 3 shows that the average lexical density of Roger Ames's translation is 57.81, nearly 3 words higher than James Legges's translation, Which means that Roger Ames uses more complex words in the translation, which might contribute the higher difficulty of reading. This support is our first assumption about the frequency of function words in the translated texts. To

make the point more clear, we have compare the complex words in each translation, and the result is presented in the following chart.

To support the statistical results, let us look at the following example:

Examples 1) 三年无改于父道, 可谓孝矣. (The Analects, 4.20)



Ames : The Master said, “A person who for three years refrains from reforming the ways of his late father can be called a filial son.”

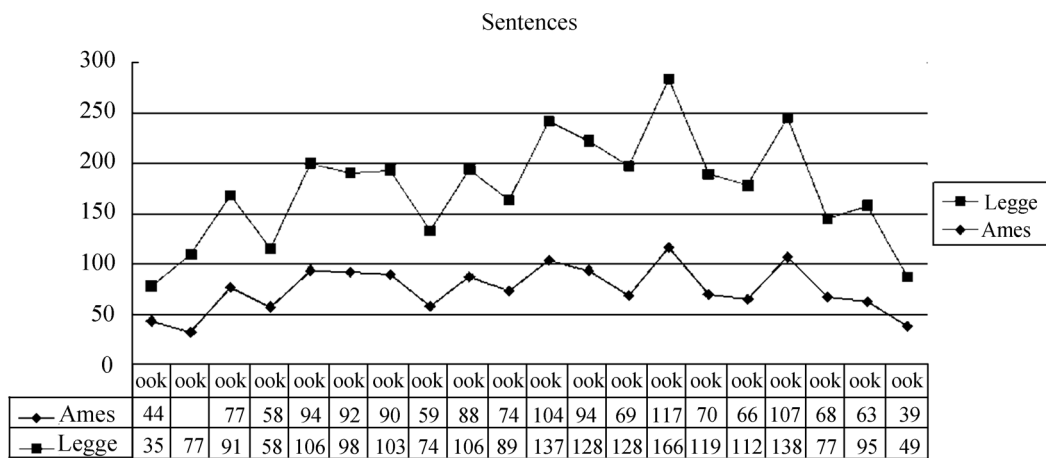
Legge: The Master said, “If the son for three years does not alter from the way of his father, he may be called filial.

From the clause structure, we can see that both Ames and Legge use compound clause, but Legge choose to use the conditional clause, which in according to the structure of the original text. But in Ames’ version, there are two content words like “refrains” and “reform”, which contributes the higher lexical density of the sentence. As far as the semantic difference, Ames’ translation implies that the son must have fully understand the ways of his

father then reform them and make them appropriate to his own particular circumstances. From the above example, we can see that the translation of Roger Ames is complex than Legge’s and there are more complex words in Ames’s translated text. Lexical density has been proved to be a useful tool to measure the amount of information and size of a text by various researchers, and our findings seem to in favor of the theory concerning about word numbers and complexity.

### 3.2.2 Sentence Simplification

Meanwhile, sentence is another variable in translation texts that indicates the simplification tendency. The following Graphs are the statistics about sentence number and word per sentence we have summarized.



**Figure 4**  
**Sentence Number of Two Translations**

As far as the number of sentences in English translations is concerned, we have found that there are two characteristics occurred. Firstly, both the sentence numbers are a little bit higher than the original text. We own this to the typological difference between English and Chinese, as we all known that as a language depends less on grammatical function words like connectives, prepositions and other types of empty words, while Chinese is indicated more paratactic characteristics rather than hypothetic features, which can be spotted in words and clauses that implicated grammatical meaning and logical relationship. The other feature we have found is that the number of sentence is to some extent, in accordance with the readability level. To be more specific, if the readability level is lower, the sentence number tends to be small, and vice versa. In the previous section, we have proved that the readability figure of Ame’s English translation is lower than Jegge’s, and now we have found that the sentence number of Roger Ames is also less than Legge, which complies our hypothesis about the relevance of readability level and sentence simplification strategy. Translators used to resort longer sentences to make the same meaning more

explicit, especially in Chinese to English translation. Therefore, Legge, a famous sinologist in great Britain, is a missionary from the London Commission. His purpose of translating the Chinese classics is to make the Chinese to believe in God, just like the Westerners. So, simple and readable are the first criteria for translation. Roger Ames and Rosemong are philosophers and have had training in formal linguistics and in Chinese philosophy. Their personal temperament is reflected in translating in the Analects and their emphasis on the philosophical significance of this Confucian classic make the English translation more like a westerner’s philosophical coursebook rather than an ancient Chinese dialogical. Let us consider the following example.

Example 2) 子曰：兴于《诗》，立于礼，成于乐。  
 (The Analect 8.8)

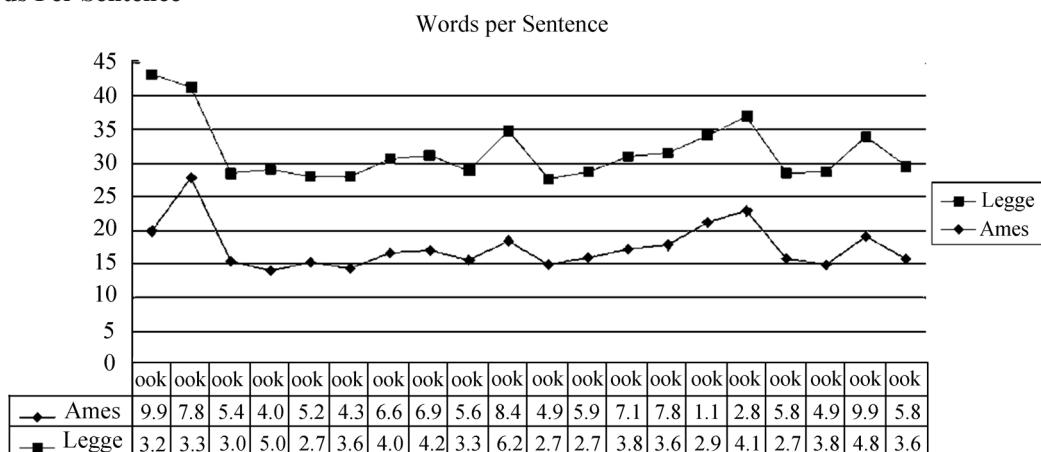
Legge: The Master said, “ It is by the Odes that the mind is aroused. It is by the Rules of Propriety that the character is established. It is from Music that the finish is received.

Ames: The Master said, “ I find inspiration by intoning the songs; I learn where to stand from observing ritual propriety (Li 礼) and I find fulfillment in playing music.

From the above English translations, we can see that Ames & Rosemont highly appreciate the significance of the key Confucian terms. When translation, they are alert to the great difference between the two cultures and guard against adopting terms saturated with Western connotations to interpret Chinese ones trying to give the true face of the original text by keep the original Chinese characters in the translation, which make the sentences

more complex and longer. While Legge's primary purposes are to introduce native Chinese culture and oriental philosophy to the Christen missionary so that they can know more about China. Thereby he tries to keep the original features and translates it literally, which makes his version word-for-word or even hard to understand, but quite concise and shorter, reserving the Ancient Chinese characteristics.

### 3.2.3 Words Per Sentence



**Figure 5**  
**Word Per Sentence of Two Translations**

The line in Figure 5 shows that the average words number in a sentence of two translations is slightly varied, especially in book 1 and book 2. Generally speaking, Ames's average sentence length of each book is higher than Legge's translation, except book three, which is 12 words. It also can be found that the sentence length of each book in Ames's translation fluctuates greater than Legge's.

Example 3) 子曰：“君子有三畏：畏天命，畏大人，畏圣人言。小人不知天命而不畏也，狎大人，侮圣人言。” (The Analect 16.8)

James Legge: There are three things of which the superior man stands in awe. He stands in awe of the ordinances of Heaven. He stands in awe of great men. He stands in awe of the words of sages.

The mean man does not know the ordinances of Heaven, and consequently does not stand in awe of them. He is disrespectful to great men. He makes sport of the words of sages.

Ames: Confucius said, Exemplary persons hold three things in awe: The propensities of tart, persons in high station, and the words of the sages persons, knowing nothing of the propensities of tian do not hold it in awe; they are unduly familiar with persons in high station, and ridicule the words of the sages.

In the English translation of James Legge, there are eight sentences. And most of them are simple sentences with the 6-9 words in a sentence; While the number of sentences in Roger Ames's translation is less and the structure of the sentences is more complex. In terms of the sentence length,

Ames's translation is longer than James Legge, with 22-26 words in a sentence. In other words, Ames tends to use long and complex sentence in his translation, while Legge prefer to use short and simple sentence.

### 3.2.4 Summary of Statistic Findings

**Table 3**  
**Comparison of Average Value**

Average value for	Roger ames	James legge
Chars per word	4.4965	4.333
Syllables per wor	1.3605	1.285
Words per sentence	17.5075	14.2055

**Table 4**  
**Comparison of Total Number**

Total number for	Roger ames	James legge
Characters	145,026	152,206
Non-space characters	118,704	122,912
Letters/numbers	114,929	118,259
Words	25,634	27,396
Complex words	2,209	2,073
Syllables	34,779	35,076
Sentences	1,506	1,986

In the above passages, we have applied the readability statistics to explain the simplification tendency in translations from two aspects: lexical density, and sentence numbers as well as words per sentence. In terms

of lexical density, translation of Roger Ames is higher than James Legge, and the average value for words per sentence is also higher than James Legge, which indicates that words and sentences in Ames's translation will be more complex than James Legge. As for the sentence, the number of Jegge is higher than Roger Ames, and we believe this is due to Legge's using more simple sentences rather than compound and complex sentences.

## CONCLUSION

As a method of assessing information of text, readability analysis used to be confined in the research areas like language teaching, textbook analysis, and reading comprehension analysis. This research however, applied readability theory into translation studies of *Lun Yu* or "the *Analects of Confucius*" in Chinese and English versions.

The analysis explains the readability statistics in two ways: the reading ease score and reading level index by using the readability formulas. Our findings show that the readability of Roger Ames's translated texts is higher than James Legge's; hence is much more difficult to read. In order to find out the degree of simplification in the translations, we also compared the two translated texts in terms of four-readability index: word numbers, lexical density, and sentence numbers and average sentence length. It is proved that, with the help of some corpus tool, readability index such as content words number, sentence length, words per sentence, can be fully used to explain certain aspects of simplification feature in translated text.

Limitation of the study: readability formulas typically are designed and used to estimate the difficulty of two surface-level features (i.e., vocabulary and syntax) of language and reading in paragraph text form. The formulas do not consider context features like cultural information or text register, or genre) that also influence text understanding.

## REFERENCES

- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp.223-250). Philadelphia / Amsterdam: John Benjamins.
- Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223-243.
- Baker, M. (1996). Corpus-based translation studies: The challenges that lie ahead. In H. Somers (Ed.), *Terminology, LSP and translation: Studies in language engineering in honor of Juan C. Sager* (pp.175-186). Amsterdam & Philadelphia: John Benjamins.
- Blum-Kulka, S., & Levenston, E. A. (1983). Universals of lexical simplification. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp.119-139). London: Longman.
- Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, 26(1), 19--26.
- Drury, A. (1985). Evaluating readability. *Professional Communication, IEEE Transactions On*, (4), 11-14.
- DuBay, W. H. (2004). *The principles of readability*. Impact Information, California.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221.
- Feng, Q. X. (2006). *Readability analysis of Burton Watson's English version of the Tso Chuan* (Unpublished MA dissertation). Dalian University of Technology.
- Guo, W. H. (2010). *Research on readability formula of Chinese text for foreign students* (Unpublished MA dissertation). Shanghai Jiaotong University.
- Huang, G. W. (2012/2013). Analyzing the reporting clause in translating confucius's *Lun Yu* (The Analect). In Y. Fang & J. J. Webster (Eds.), *Developing systemic functional linguistics: Theory and application*. London: Equinox.
- Jia, C. Y. (2004). *Ying Han Xin Wen Yi Du Xing Bi Jiao Yan Juan* (Unpublished MA dissertation). Shanghai Foreign Language University.
- Klar, K. E., (1963). *The measurement of readability* (p.1). Ames, IA, Iowa State University Press.
- Laviosa, S. (2002). *Corpus-based translation studies: Theory, findings and applications*. Amsterdam: Rodopi.
- Toury, G. (1995). *Descriptive translation studies—And beyond*. Amsterdam and Philadelphia, PA: John Benjamins.
- Venuti, L., & Baker, M. (1998). *Routledge encyclopedia of translation studies*. London: Routledge.
- Vanderauwera, R. (1985). *Dutch novels translated into English: The transformation of a "minority" literature*. Amsterdam: Rodopi.
- Wang, X. M., & Yang, Y. J. (2012). Research and design of English textbook evaluation based on readability and FAHP. *Foreign Languages in China*, (03), 67-77.
- Wang, Z. P. (2012). An appraisal of translation difficulty. *Translations*, (06), 153-167.
- Xiao, Y. Y. (2012). *Jiao Ke Shu Ke Du Xing Ping Jia Yan Jiu*. Jinan, China: Shandong Normal University.
- Xing, F. K., & Cheng, D. Y. (2010). English text readability based on statistic models. *Journal of PLA University of Foreign Languages*, (11), 19-24.
- Yang, J. Y. (2008). *The study on measurement of language difficulty of advanced Chinese intensive reading textbooks*. Beijing University.
- Zhang, C. (2010). Research on English readability: Retrospection and reflection. *Journal of Hunan Institute of Engineering*, (3), 47-51.
- Zamaniah, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory and Practice in Language Studies*, 2(1), 43-53.
- Zhou, L. L. (2012). *A case study of improving readability by reducing translation of English prepositions articles and pronouns into Chinese* (Unpublished MA dissertation). Shanghai, China: Shanghai Foreign Language University.