Investigating the Simplification Hypothesis in Persian Translated, Authored, and Pseudotranslated Texts: A Corpus-based Analysis of Zabiollah Mansouri's Works

Masoud Sharififar

Shahid Bahonar University of Kerman, Kerman, IRAN

Mina Zandrahimi

Shahid Bahonar University of Kerman, Kerman, IRAN

Asma Bahadori

Shahid Bahonar University of Kerman, Kerman, IRAN

ABSTRACT

Laviosa's simplification hypothesis is a subcategory of translation universals, suggesting that translated texts are of lower lexical variety. The current study investigated Laviosa's simplification hypothesis in Persian translated, authored, and pseudotranslated texts. To that end, three books by Zabiollah Mansouri were selected; a three-fold comparative analysis was conducted, containing 15,507 words of translated texts, 15,396 of authored texts, and 15,607 words of pseudotranslated texts. The corpus analysis tool AntConc 3.5.7 was leveraged to count types, tokens, and lexical density. Then ANOVA and Dunnett Post Hoc tests were conducted to analyze the data, and results revealed that in terms of type/token ratio, there was no significant difference between translated, authored, and pseudotranslated texts. The lexical density of authored texts was significantly higher than that of translated and pseudotranslated texts, but there was no significant difference between the lexical density of translated and pseudotranslated texts. Regarding mean sentence length, authored texts presented a significantly higher mean sentence length than translated and pseudotranslated texts; besides, the mean sentence length of pseudotranslated texts was significantly higher than that of translated texts. The results of the study revealed that Mansouri succeeded in convincing readers to accept his work as a translated text.

KEYWORDS: lexical density, sentence length, simplification hypothesis, translation universals hypothesis, type/token ratio

1. Introduction

In the last decades, the field of Translation Studies has experienced an upsurge of interest in descriptive approaches, focusing on the investigation of translated texts to mark out regularities and norms that characterize the language of translations (Piccioni, 2013). For instance, corpus-based research in translation studies proposes that translated texts display some distinctive linguistic features, termed "translation universals" by Baker (1993), such as a general tendency toward simplification and explicitation, which are believed to occur independently of the specific language pairs engaged in the process of translation (243). According to Chesterman (2004b:39), universals of translation are claims about the common properties of translated texts which can be either S-universals: "universal differences between translations and their source texts" or T-universals: "universal differences between translations and comparable authored texts, i.e., characteristics of the way translators use the target language".

While describing translation features, researchers confronted some translations with no source texts. According to Popovič (1976:20), "an author may publish his original works as a translation in order to win a wide public, thus making use of the readers' expectation". For various reasons such as ideology of patronage, fear of censorial measures and specific cultural and literary plans, authors may present their original texts as translations (Sharififar et al., 2018). Popovič (1976) refers to such texts as pseudotranslations (as cited in O'Sullivan, 2010:123).

According to Toury (1995), pseudotranslations mirror standard features of translations since pseudotranslators apply linguistic and textual items common in translations (as cited in Shuttleworth & Cowie, 1997). Pseudotranslators apply linguistic characteristics of translations to their pseudotranslated texts to convince the readers that they are genuine translations. Therefore, despite the lack of actual source texts, pseudotranslations have been discussed within the framework of descriptive translation universals (Rambelli, 2009; Toury, 2012). In Rizzi's (2008:154) view, the relevance of pseudotranslations to Translation Studies is undeniable; nevertheless, "the nature and function of pseudotranslations are still far from clear".

1.1. Statement of the Problem

Chesterman (2004a:39) points out that translation universals are only hypotheses (when testing through empirical evidence). Some have been substantiated more than others, and some evidence rejected the hypotheses. Thus, "in most cases, the jury is still out" (ibid.). Accordingly, more research on a wider variety of data would contribute to determining the scope of these hypotheses and may eventually lead to the formulation of further theorization in this area. The hypothesized universals of translation in general, and simplification in particular, have focused on European languages (Chesterman, 2010). The goal of the present study is to test whether the simplification hypothesis can be applied to less-explored languages such as Persian. Indeed, further research could reveal which languages support a given universal tendency.

Laviosa (2002:43-51), whose simplification hypothesis has been adopted as the theoretical framework of this article, contends that the evidence she gathered for the simplification hypothesis in the English comparable corpus may not apply to all text types and for all given parameters. Furthermore, an area neglected by most translation scholars is the applicability of translation universals, such as simplification, to pseudotranslations. When a writer publishes his/her work as a translation of someone else's work, it is a less-explored area whether his/her language is similar to the language of translators or not, and it is unknown whether the linguistic characteristics of pseudotranslations are similar to translations or not. In other words, the researchers of the present study seek to know if the author applied linguistic features of translations to his pseudotranslation to convince readers that it is a genuine translation.

1.2. Objective of the Study

The present inquiry investigates the applicability of Laviosa's (2002) simplification, as a hypothesized translation universal, to the selected Persian translated texts. Moreover, the researchers hope to explore a new perspective: the applicability of the simplification hypothesis to pseudotranslations. In other words, the goal of the present study is to determine whether there are significant differences between translated, authored, and pseudotranslated

texts regarding the frequency of simplification features in Mansouri's (1984, 1985) works. This study also attempts to provide the answers to the following research questions.

1.3. Research Questions

1. To what extent has the simplification hypothesis been applied in the selected Persian translated, authored and pseudotranslated texts?
2. Does the probable applicability of simplification hypothesis in selected Persian translated, authored and pseudo translated texts verify the simplification hypothesis?

1.4. Significance of the Study

According to Chesterman (2004a:33): "Any science seeks generalities. The aim is to transcend knowledge of particular cases by discovering general regularities or laws or general descriptive hypotheses that cover more than a single case". He asserts that science can make predictions about future or unstudied cases by investigating similarities between single cases and then generalizing based on these cases. Accordingly, the present research will report on a case study that sheds new light on the applicability of the simplification hypothesis, specifically to Persian translated and pseudotranslated texts.

The concept of T-universals is an attractive object of study, given its comparative and contrastive nature, as it provides the possibility of comparing translations of different source texts not with their source texts but with each other. T-universals in general and simplification, in particular, are a subject that keeps the door open for further research. In Iran, to the best of the researchers' knowledge, it is far from having been comprehensively explored.

Since translation hypotheses have usually been centered around European languages, more research on other less-explored languages such as Persian may contribute to the existing body of knowledge within Translation Studies. Moreover, regarding the simplification hypothesis, Laviosa (2002:77) contends that the evidence she reported for the English Comparable Corpus "is by no means uniform in all texts analyzed and for all the parameters considered". She argues that further studies regarding translation universals may "discover more and more

facets of the nature of translated texts and translating and raises awareness about a complex, reciprocal relationship that links language to culture".

To the best of the researchers' knowledge, limited work has been done on the applicability of the simplification hypothesis to pseudotranslated texts. The present research addresses a relatively underexplored area in translation studies. Accordingly, it sheds light on the characteristics of pseudotranslated texts, when they are probed within the framework of translation universals.

1.5. Limitations of the Study

This study faced two main limitations. First, Zabiollah Mansouri (1895-1986), whose works are used in this research, was a well-known translator according to people of his time. He was a prolific translator of historical books, and the breadth of his works has given a prominent position among his contemporaries. It is worth mentioning that the dates of his life were taken from a biographical reference book by Salehyar (1972: 135). The reason for choosing Mansouri in this research lies in the fact that his works were historical books and thus of interest to his readership. Second, some of his books are not accessible in their original forms, because they were published before the Islamic Revolution in Iran, and the available ones may have undergone censorship by controlling entities.

As far as the researchers speculated, Mansouri published his works as translations not only to win popularity among his readers but also to state his opinion under the name of more well-known writers and waive up any responsibility concerning its consequences. And according to Sharififar et al. (2018), somewhere, Mansouri referred to some of his translations as "adaptations". Therefore, his mastery of the target language rather than the source language may influence the results.

2. Translation Universals: A Review

According to Laviosa (as cited in Baker & Saldanha, 2011), the main universals of translation are explicitation, normalization, and simplification. The present research investigates the application of the simplification hypothesis in the selected pseudotranslations.

2.1. Simplification

Baker (1996) claims that translators tend to simplify the language of translations. According to Ghadessy and Gao (2001), simplification is considered a conscious strategy in foreign language learning to prepare instructional materials. For example, reading texts are consciously simplified by reducing the complexity of vocabulary and sentence structure.

At the word level, shorter words are considered more straightforward than longer ones; concrete words are simpler than abstract ones; function words are considered more straightforward than content ones; high-frequency words are considered simpler than low-frequency ones. At a higher level, single-item verb and noun phrases (i.e., verb and noun phrases with a single head) are more straightforward than multiple-item ones. At the sentence/clause level, translators subconsciously choose more straightforward sentences than compound and complex ones (Ghadessy & Gao, 2001).

Laviosa's (2002) research on a comparable corpus of English newspaper articles and narrative prose revealed that translated texts were of lower lexical density, type/token ratio (lexical variety), and mean sentence length, when compared to original texts.

2.1.1. Lexical Density

Lexical density, as Johansson (2009:65) puts it, is the "proportion of content words (nouns, verbs, adjectives, and often also adverbs) to the total number of words". It is closely related to the notion of information packing or information load, i.e., "a text with a high proportion of content words contains more information than a text with a high proportion of function words (prepositions, interjections, pronouns, conjunctions, and count words)" (ibid.).

The notion of lexical density was proposed for the first time by Halliday (1985/1994) as an index that differentiates spoken language from written. It is believed that texts with a high degree of lexical density are more difficult than those with a lesser degree of lexical density, resulting in written texts with complex grammar (as cited in Ghadessy & Gao, 2001). Baker (1995) states two reasons for higher lexical density and, as a result, the greater difficulty of written texts compared to spoken texts. The first reason behind such difficulty is that written

texts are context-free compared to spoken ones, in the immediate contexts that make them simpler. And the second reason is that written texts are highly edited. Therefore, redundancy and repetition are reduced.

Different methods have been suggested for measuring the lexical density of a text. According to Chafe and Danielwicz (1987:5), two methods can be leveraged to measure lexical density: 1) measuring the ratio of content words to all words of the text; and 2) measuring the ratio of content words to ranking clauses (not embedded) of a text. Halliday (1985/1994:351-352) suggested that lexical density can be measured by dividing lexical items by the number of ranking clauses (not embedded). Baker (1995) defines lexical density as "the percentage of lexical as opposed to grammatical words in a specific text, and it can be measured by dividing the number of lexical words by the total number of words in a text, and the result multiplied by 100 to arrive at the percentage" (237).

2.1.2. Lexical Variety

Lexical variety, or lexical diversity, can be measured via the calculation of type/token ratio (Kruger, 2002). Types refer to "the number of different words in a text" and tokens are "the total number of running words of a text (Johansson, 2009:62). A low ratio of types to tokens implies less varied vocabulary and, as Laviosa (2002) suggests, a simpler text.

2.1.3. Mean Sentence Length

Some scholars noticed that translations are simplified stylistically. For instance, Malmkjaer (2004) points out that punctuation is stronger in translated texts. Full stops usually take the place of semicolons, and semicolons or full stops commonly take the place of commas. Accordingly, complex and long sentences in the source texts are simpler and shorter in the target language. Hence, translated texts are of shorter sentence length compared to original texts.

2.2. Pseudotranslation

In line with Popovič, Toury (1980) defines pseudotranslations as "target language texts which are regarded in the target culture as translation though no genuine source texts exist for them" (as cited in Shuttleworth & Cowie, 1997:134). To Toury (1980:83), pseudotranslations are

important for Translation Studies for two reasons. First, they "introduce innovations into a literary system, especially when it is resistant to deviation from canonical models and norms". Second, they "provide a useful insight into prevailing notions of features which characterize translated texts" (as cited in Shuttleworth & Cowie, 1997:134).

According to Toury (2012:47), text producers, translators, text consumers, and overall members of a culture are aware of the position of translations in a culture and the function they may carry out. He asserts that: "[t]his awareness is often accompanied by, and finds its expression in certain behavioral patterns, including textual-linguistic features". Sometimes, one may manipulate this awareness and compose texts but then present them as translations; it is interesting to note that from the standpoint of the host culture, these pseudotranslations are genuinely on a par with accurate translations (Toury, 2012).

2.2.1. Motivations for Writing Pseudotranslation

2.2.1.1. Introducing Novelty into a Culture

Toury (1995) argues that a critical factor for producing and distributing an original work under the guise of translation is that the writer finds it an easy and trouble-free method of importing novelties to a culture. He notices that this novelty is not confined to the realm of literature (2012). This is the case, especially for those nations that are extremely loyal to their national cultures and resistant to alien cultures. The pseudotranslator "put the cultural gate keeper to sleep as if it were translated" (2012:4). The writer resorts to pseudotranslation to introduce innovations to a culture that is resistant to change, so the culture will be changed without sensitizing society, because cultural innovation in translation meets with a much greater degree of tolerance (1995). Actually, in this case, pseudotranslation can act as a culture-planning device (2012).

2.2.1.2. Fear of Censorial Measures

Another reason for disguising an original work under the veil of translation is that authors use pseudotranslations to shield their works from censorial measures: "deviations from what is culturally (or governmentally) sanctioned that are most likely to meet with opposition" (Toury 1995:41). Censorship is more lenient with translated texts, compared with authored

texts. One explanation for this difference is that the "non-domestic origin of translations makes them look less menacing; another is that there seems to be no way of actually going after the absent author, who presumably takes most of the blame" (ibid.).

2.2.1.3. Popularity of Foreign Works

The attitude of the readers is an important factor that may cause the author to publish his/her own works under the guise of translation. In such cases, foreign works enjoy greater popularity compared to works originally written in the national language. Accordingly, translated books are sold more often than those published in the originally authored language. Rizzi (2008:155) held that disguising the authorship of a pseudotranslation provided an opportunity to "conceal his/her authority or to exploit the dominant culture".

2.2.2. Pseudotranslation and Polysystem

Even-Zohar (1990:2) coins the term *polysystem* to refer to "a multiple system, a system of various systems which intersect with each other and partly overlap, using concurrently different options, yet functioning as one structured whole, whose members are interdependent". This concept provides an opportunity to describe and analyse the functions and evolution of literary systems. The relation and position of these systems change based on the historical moment and occur in a dynamic hierarchy (Munday, 2012). If at a given time, an innovatory literary type or genre occupies the highest position in a hierarchically conceived system, then the conservative literary type occupies a relatively lower position in this literary polysystem (ibid.).

Even-Zohar (1990) adds that the position of translated literature in the polysystem even determines the translation strategy. In the primary position, translators feel free to break target literature conventions and produce adequate translations. The influence of foreign literature may give rise to new models in the target language for both authored and translated literature. In the case of secondary positions, translators feel constrained to follow target literature conventions. Accordingly, they produce more non-adequate translations that conform to the target literature conventions (Munday, 2012).

3. Research Data

To conduct this study, the researchers used three books authored by a Persian writer and translator named Zabiollah Mansouri (1895-1986). His main expertise was in pseudotranslation rather than translation. He created attractive titles for his books and sold them as translations attributed to fictitious authors. According to Braheni (1985:98-99), he changed a novel of 600 pages into 100 pages. He also distorted historical facts in his translations. For example, in his translation of *The Lord of Alamoot*, he reported the death of Neẓām-al-Molk during hunting, when in fact he died on his trip to Baghdad. His work, concerning the aforementioned case, was selected for this research. The data includes one original Persian book, one Persian book as pseudotranslation, and one Persian translation of an English book. The selected books and the justifications for choosing them are elaborated in the following subsections.

3.1. A Year amongst the Persians

A Year amongst the Persians is an English travelogue written by an English orientalist named Edward Granville Browne (1893). As the title shows, this book was the outcome of Browne's one-year visit to Iran. It provides information on Iranians' social, political, and economic conditions in Nāṣer-al-Din Shah's era. The researchers of the present study have used the Persian translation of this book by Zabiollah Mansouri.

3.2 Islamic Sciences in the first Century A.H.

To the best of the researchers' knowledge, *Islamic Sciences in the first Century A.H.* (1984) is the only book on which Mansouri's name is printed as author. This book provides information on the Muslim lifestyle and Islamic sciences in the first century A.H.

3.3 The Lord of Alamoot: Hasan Sabbah

The Lord of Alamoot (1985) was claimed to be written by Pol Amir, a French author, and translated by Zabiollah Mansouri. To the best of the researchers' knowledge, there was no one named Pol Amir in the history of French literature. Accordingly, considering Toury's (1995) definition of pseudotranslation, *The Lord of Alamoot* could be regarded as a pseudotranslation.

This book describes the life of Hasan Sabbah, the founder of the Batini sect, and the social and political conditions of Iran at his time. The writer (assumed translator or pseudotranslator) stated that the Sabbah's movement was a religious movement aimed to liberate Iranians from the yoke of Abbassi's Caliphs or the local kings of Iran under their influence.

4. Data Collection and Analysis

The researchers of this study selected almost an equal number of words from the three books under study. To that end, some pages of each book have been selected randomly. It should be mentioned that the number of randomly selected pages was different, but the number of words was almost equal, i.e., 55 pages of the translated book, 46 pages of the pseudotranslated book, 47 pages of the original book. Pages were selected randomly to get a fair number of words. i.e., 15,507 words of translated texts, 15,396 of authored texts, and 15,607 words of pseudotranslated texts.

AntConc 3.5.7 software, a corpus-analysis tool, was used to count types, tokens, lexical words, and mean sentence length of the corpora. Subsequently, in order to analyse the data with SPSS software, the researchers divided the randomly selected pages of each book into five sections.

Three text groups were analysed regarding three features of the simplification hypothesis, including lexical density, type/token ratio, and mean sentence length. According to the simplification hypothesis (Laviosa, 2002), translations tend to be simpler than non-translations. To confirm the applicability of the simplification hypothesis to Persian language, translated texts must represent lower lexical density, lower mean sentence length, and lower type/token ratio, compared to authored texts. And to substantiate the applicability of the hypothesis to pseudotranslated texts, the texts must show lower lexical density, lower mean sentence length, and lower type/token ratio, compared to authored ones.

All three groups of texts were extracted from printed books. Thus, to process the texts with AntConc, it was necessary to convert the printed material into electronic format. Texts were

converted into Microsoft Office Word 2013 readable format. Next, texts were saved in Unicode format in Notepad. In line with the theoretical framework, to investigate the simplification hypothesis in three text groups, the researchers were required to measure the type/token ratio, lexical density, and mean sentence length of the selected texts. In the next stage, AntConc is leveraged to calculate the types, tokens, and lexical words of the texts under study.

4.1. Type/Token Ratio

Tokens refer to "the total number of words in a given corpus" (Williams and Chesterman, 2002:98). Concerning electronic texts, "any sequences of letters with an orthographic space on either side is counted as a word or more precisely a *token*" (Baker, 1995).

Types are defined as the total number of different words (or, as its name reveals, word types) in a given corpus (Williams & Chesterman 2010:98). Many words may be used more than once in a corpus, but for counting the number of types, each word is counted once.

To measure the type/token ratio, the researchers converted the three groups of texts into a text document format, then AntConc 3.5.7 software provided types and tokens. The researchers measured the type/token ratio with the following formula:

Type/token ratio = (number of types / number of tokens) * 100

4.2. Lexical Density

Lexical density refers to the number of content words to the total number of words (Stubbs, 1996). To measure lexical density, the researchers used the following formula:

Lexical density = (number of lexical words / total number of words, or tokens) * 100

4.2.1. Mean Sentence Length

Mean sentence length refers to the "average number of words per sentence in a text" (Malmkjaer, 2004:4). To measure mean sentence length, the following formula was used:

Mean sentence length = number of words in the text / number of sentences in the text

The number of sentences was counted manually, and the number of words (tokens) was obtained through AntConc software.

4.2.2. Corpus-based and Statistical Analyses

In order to test the applicability of the simplification hypothesis to the selected Persian translated and pseudotranslated texts, the three groups of respective texts were compared and contrasted in terms of three features of the simplification hypothesis. Each text group was investigated for the above features with AntConc.

After measuring simplification features, i.e., type/token ratio, lexical density, and mean sentence length, the raw data were analyzed with SPSS software for statistical analysis to calculate each simplification feature's mean score and standard deviation. An ANOVA test was conducted to determine whether the observed difference between means was statistically significant or not. When ANOVA presents a significant difference between means, it does not report which pairs of means are different. Therefore, the Dunnett Post Hoc test was conducted to identify the pairs with a significant difference. Besides, Shapiro-Wilk and Levene tests were applied to test the normality and homogeneity of the data.

The collected data regarding simplification features (type/token ratio, lexical density, and mean sentence length) of translated, authored, and pseudotranslated texts is presented in tables, and the results are compared with each other.

4.3. Results

The present section is devoted to the results, data analysis, and discussion of the results. Simplification features of the three text groups have been compared through statistical tables. As previously mentioned, simplification features can be observed with type/token ratio, lexical density, and mean sentence length.

4.3.1. Type/Token Ratio

The types and tokens of each text group were measured with AntConc 3.5.7 software. The observed type/token ratios of translated, authored, and pseudotranslated texts are presented in Table 1.

Table 1: Type/token ratio calculation for the three text groups

Text type	Sections	Type/Token ratio
Translated	1	34.58
	2	33.99
	3	32.19
	4	34.58
	5	32.09
Authored	1	33.98
	2	31.14
	3	33.2
	4	33.95
	5	30.41
Pseudotranslated	1	32.6
	2	26.94
	3	31.81
	4	32.46
	5	33.81

Then, numerical results from AntConc were processed with SPSS to calculate the mean type/token ratio of translations, authored texts, and pseudotranslations. Table 2 features the resulting mean scores for the three text groups. As mentioned earlier, to analyse the data with SPSS software, the researchers divided the randomly selected pages of a book into five sections. Therefore, N stands for the number of each text group's sections.

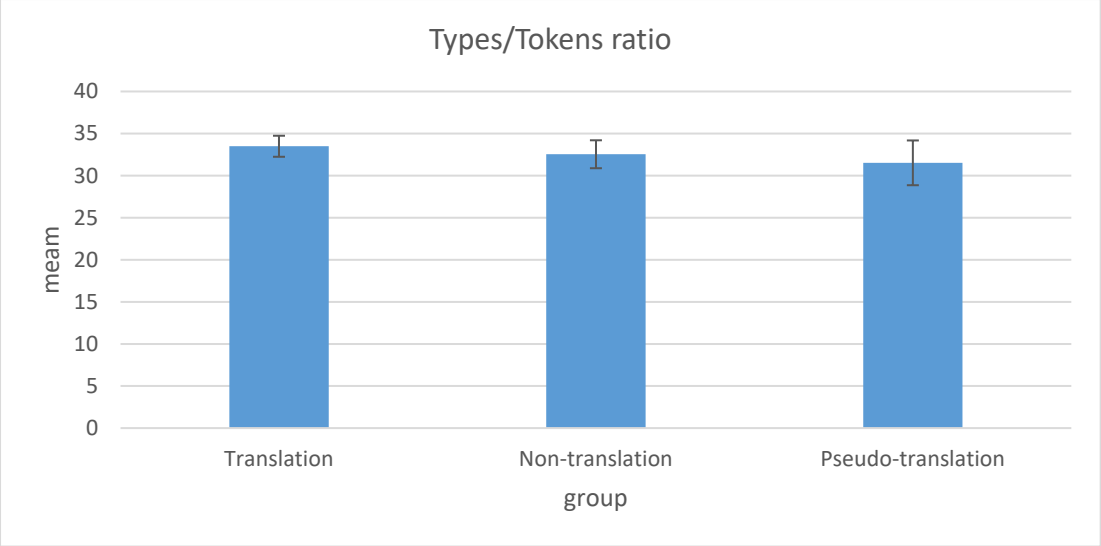
Table 2: Descriptive statistics for type/token ratio in the three text groups

Group	N	Minimum	Maximum	Mean	Std. Deviation
Translated	5	32.09	34.58	33.49	1.25
Authored	5	30.41	33.98	32.54	1.66
Pseudotranslated	5	26.94	33.81	31.52	2.66

As the statistical results show, the mean lexical density of translated texts was 33.49 ± 1.25, the mean lexical density of authored texts was 32.54 ± 1.66, and the mean lexical density of pseudotranslated texts was 31.52 ± 2.66.

The mean type/token ratio of translated, authored, and pseudotranslated texts is displayed on a bar graph in Figure 1.

Figure 1: Mean (±Std. deviation) for type/token ratio in the three groups



4.3.2. Checking Statistical Assumptions

Parametric tests require a set of assumptions. Normality of variances is considered "one of the essential assumptions for drawing reliable inferences about the underlying population of data" (Verma & Abdel-Salam, 2019:66). To this end, the normality of the observed data has been checked through Shapiro-Wilk test. Besides, the assumption of homogeneity of

variances between groups was analysed through Levene test. Results are presented in the following tables.

To determine whether the data concerning type/token ratio meet the normality assumption, Shapiro-Wilk test was conducted. The output of Shapiro-Wilk test is shown in Table 3.

Table 3: Results of Shapiro-Wilk test for type/token ratio

	Group	Shapiro-Wilk		
		Statistic	Df	Sig.
Type/token	Translated	.795	5	.073
	Authored	.846	5	.182
	Pseudotranslated	.798	5	.079

The significance value (p-value) of the Shapiro-Wilk test is more than 0.05. Therefore, the normality assumption of the observed data has been met, i.e., the observed data concerning type/token ratio of translated, authored, and pseudotranslated texts are normally distributed. Furthermore, Levene's test was conducted for homogeneity of variance, and the output is reported in Table 4.

Table 4: Result of Levene test for type/token ratio

	Levene statistic	Df1	Df2	Sig.
Type/Token	.660	2	12	.535

As observed, the assumption of homogeneity of variances between three groups was met (p=0.535).

ANOVA (Analysis of Variance) test was then conducted to determine whether there exists a significant difference among three or more means (Verma & Abdel-Salam, 2019). To detect whether the difference between means was statistically significant, a one-way ANOVA test was conducted using SPSS for the three obtained mean scores of type/token ratio. Table 5 indicates the results.

Table 5: Results of ANOVA test for type/token ratio

	Sum of Squares	Df	Mean /Square	F	Sig.
Between Groups	9.636	2	4.818	1.266	.317
Within Groups	45.653	12	3.804		
Total	55.288	14			

As can be inferred from the above table, there was no statistically significant difference between the mean type/token ratio of translated, authored, and pseudotranslated texts (p=0.317).

The standard threshold for statistical significance is $p \leq 0.05$. Since 0.317 is greater than 0.05, we fail to reject the null hypothesis. The null hypothesis for an ANOVA states that all group means are equal. The high p-value (0.317) indicates that the observed differences in the mean type/token ratios between the translated, authored, and pseudotranslated texts are not statistically significant and are likely due to random chance.

4.3.3. Lexical Density

As stated in chapter three, the lexical density of each test group was measured by dividing the number of lexical words by the tokens (the total number of words). The observed data are presented in Table 6.

Table 6: Lexical density calculations for the three text groups

Text type	Sections	Lexical Density
Translated	1	76.76
	2	77.16
	3	76.76
	4	76.58
	5	76.99
Authored	1	80.65
	2	79.56

	3	80.63
	4	80.15
	5	79.86
Pseudotranslated	1	79.06
	2	78.15
	3	75.56
	4	78.87
	5	77.38

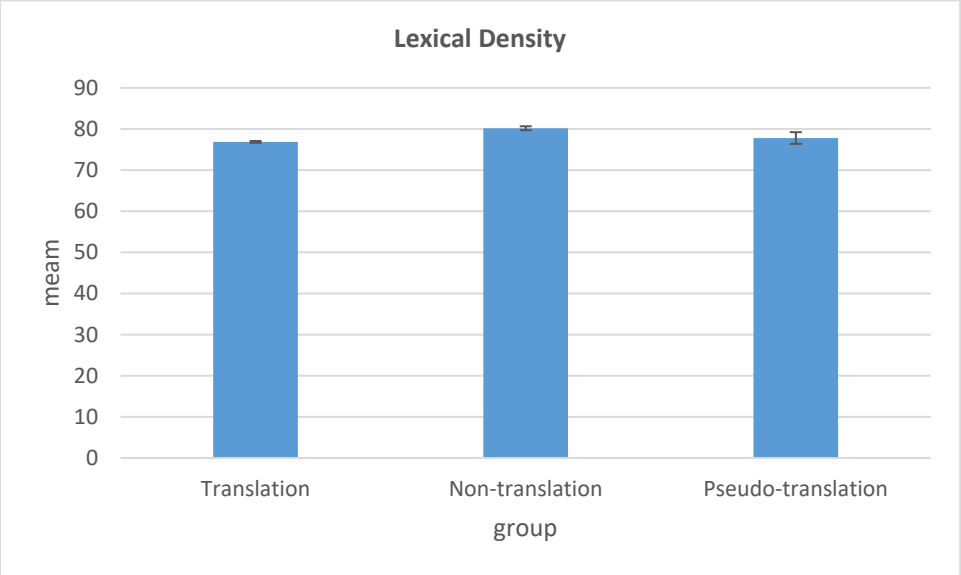
SPSS was leveraged to process the data in order to calculate the mean lexical density for translated, authored, and pseudotranslated texts. Table 7 shows the results.

Table 7: Descriptive statistics for lexical density in the three text groups

Group	N	Minimum	Maximum	Mean	Std. Deviation
Translated text	5	76.58	77.16	76.85	.22
Authored text	5	79.56	80.65	80.17	.48
Pseudotranslated text	5	75.56	79.06	77.80	1.42

As Table 7 shows, the mean lexical density of translated texts was 76.85± 0.22, the mean lexical density of authored texts was 80.17± 0.84, and the mean lexical density of pseudotranslated texts was 77.80± 1.42. The mean lexical density of the three groups is featured on a bar graph in Figure 2.

Figure 2: Mean (±Std. deviation) for lexical density in the three groups



The normality of data distribution was analysed via Shapiro-Wilk test. The results are presented in Table 8.

Table 8: Result of Shapiro-Wilk Test for Lexical Density

	Group	Shapiro-Wilk		
		Statistic	Df	Sig.
Lexical density	Translated text	.951	5	.746
	Authored text	.912	5	.478
	Pseudotranslated text	.896	5	.389

As observed, the result of the Shapiro-Wilk test shows that $p > 0.05$; therefore, the normality assumption of the observed data for translated, authored, and pseudotranslated texts was met.

To measure the homogeneity of variance of the lexical density of translated, authored, and pseudotranslated texts, Levene's test was conducted, and the result is presented in Table 9.

Table 9: Result of Levene test for lexical density

	Levene Statistic	Df1	Df2	Sig.
Lexical density	5.030	2	12	.026

The assumption of homogeneity of variances between the three groups was violated (p=0.026).

Then one-way ANOVA test was run to determine whether there was a statistically significant difference between the mean lexical density of translated, authored, and pseudotranslated texts. The output is presented in Table 10.

Table 10: Result of one-way ANOVA test for lexical density

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	28.365	2	14.183	15.979	<.001
Within Groups	10.651	12	.888		
Total	39.017	14			

As observed, the difference between the mean lexical density of translated, authored, and pseudotranslated texts was statistically significant (p< 0.001).

The ANOVA test did not report which pairs of means were different. Therefore, Dunnett Post Hoc test was run to identify the pairs with significant differences. One fixed control group was compared to all the other groups. Table 11 presents the output of the Dunnett Post Hoc test.

Table 11: Result of Dunnett Post Hoc test for lexical density

	(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.
Dunnett T3	Translated	Authored	-3.317*	.236	<.001
		Pseudotranslated	-.951	.642	.452
	Authored	Translated	3.317*	.236	<.001
		Pseudotranslated	2.366*	.669	.045
	Pseudotranslated	Translated	.951	.642	.452
		Authored	-2.366*	.669	.045

As Table 11 reports, the mean lexical density of authored texts was significantly higher than the mean lexical density of translated texts ($p< 0.001$) and pseudotranslated texts ($p=0.045$). However, the difference between the mean lexical density of translated and pseudotranslated texts was not statistically significant ($p=0.452$).

4.3.4. Mean Sentence Length

The mean sentence length was calculated by dividing the number of words by the number of sentences in each text group. Table 12 reports the results.

Table 12: Sentence length calculations for the three groups.

Text type	Sections	Sentence length
Translated	1	29.9
	2	29.28
	3	28.61
	4	26.03
	5	25.24
Authored	1	46.03
	2	63.28
	3	61.1
	4	56.08
	5	48.05
Pseudotranslated	1	38.94
	2	33
	3	38.95
	4	41.05
	5	40.88

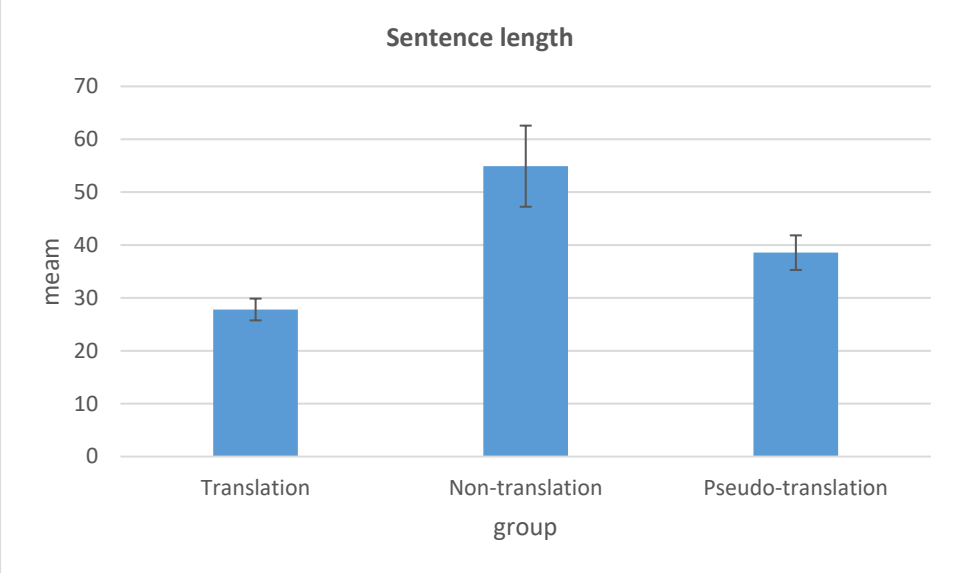
Next, data were processed with SPSS to calculate the mean sentence length for translated, authored, and pseudotranslated texts. The results are presented in Table 13.

Table 13: Descriptive statistics for sentence length in the three groups

Group	N	Minimum	Maximum	Mean	Std. Deviation
Translated text	5	25.24	29.90	27.81	2.06
Authored text	5	46.03	63.28	54.91	7.67
Pseudotranslated text	5	33.00	41.05	38.56	3.27

As Table 13 indicates, the mean sentence length of translated texts was 27.81± 2.06, the mean sentence length of authored texts was 54.91±7.67, and the mean sentence length of pseudotranslated texts was 38.56± 3.27. Results are displayed in a bar graph in Figure 3.

Figure 3: Mean (±Std. deviation) for sentence length in the three groups



The normality of the data distribution was tested by conducting a Shapiro-Wilk test. Results are reported in table 14.

Table 14: Results of Shapiro-Wilk test for sentence length

	Group	Shapiro-Wilk		
		Statistic	Df	Sig.
Sentence length	Translated text	.889	5	.351
	Authored text	.906	5	.445
	Pseudotranslated text	.792	5	.070

As the result of the Shapiro-Wilk test shows ($p > 0.05$), the normality assumption of the observed data regarding sentence length for translated, authored, and pseudotranslated texts was met.

To measure the homogeneity of variance of sentence length of translated, authored and pseudotranslated texts, Levene's test was conducted. Results are presented in Table 15.

Table 15: Result of Levene test for sentence length

	Levene Statistic	Df1	Df2	Sig.
Sentence length	6.536	2	12	.012

As observed, the assumption of the homogeneity of variances in the three text groups was not met ($p=0.012$).

To determine whether the difference between means was statistically significant, a one-way ANOVA test was run in SPSS for the three obtained means of sentence length. Table 16. indicates the results.

Table 16: Result of ANOVA for sentence length

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	1861.444	2	930.722	37.814	<.001
Within Groups	295.355	12	24.613		
Total	2156.799	14			

As observed, the difference between the mean sentence length of translated, authored, and pseudotranslated texts was statistically significant ($p< 0.001$).

Since the ANOVA test did not report which pairs of means were different, Dunnett Post Hoc test was conducted to identify the pairs with a significant difference; one fixed control group was compared to all other groups. Table 17 presents the output of the Dunnett Post Hoc test.

Table 17: Result of Dunnett Post Hoc test for sentence length

	(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.
Dunnett T3	Translated text	Authored text	-27.095*	3.554	.002
		Pseudotranslated text	-10.751*	1.728	.001
	Authored text	Translated text	27.095*	3.554	.002
		Pseudotranslated text	16.344*	3.731	.016
	Pseudotranslated text	Translated text	10.751*	1.728	.001
		Authored text	-16.344*	3.731	.016

As Table 17 shows, the mean sentence length of authored texts was significantly higher than the mean sentence length of translated texts ($p=0.002$) and pseudotranslated texts ($p=0.016$). Besides, the mean sentence length of pseudotranslated texts was significantly higher than translated texts ($p=0.001$).

4.4. Discussion

The findings of the present study confirm Laviosa's (2002) simplification hypothesis in terms of sentence length and lexical density. However, the results did not support the simplification hypothesis in terms of type/token ratio. Regarding type/token ratio, there was no significant difference between translated, authored, and pseudotranslated texts. Regarding lexical density, authored texts showed higher lexical density than translated and pseudotranslated texts. However, as can be observed in Table 4.8 and Figure 4.2, there was no significant

difference between translated and pseudotranslated texts. Accordingly, they were of the same level of simplicity in terms of lexical density. The sentence length of authored texts was significantly higher than that of translated and pseudotranslated ones, and the sentence length of pseudotranslated texts was significantly higher than that of translated texts. Hence, like translated texts, the applicability of Laviosa's (2002) simplification hypothesis to pseudotranslated texts was proved in terms of lexical density and sentence length because both pseudotranslated and translated texts registered lower lexical density and mean sentence length than authored texts. However, they did not support the simplification hypothesis in terms of type/token ratio.

It is interesting to note that in terms of lexical density and sentence length, pseudotranslated texts showed a closer affinity to translated texts in comparison to authored texts. It can be thus assumed that Mansouri was successful in writing his pseudotranslated texts in a way to resemble translated texts to conform to the readers' expectations and thereby accomplish a wider readership.

5. Answers to the Research Questions

Regarding the first research question, it should be mentioned that Laviosa's (2002) simplification hypothesis proposes that translated texts are of lower type/token ratio, lexical density, and sentence length. As far as the first feature is concerned, the research findings revealed that the difference among the type/token ratio of translated, authored, and pseudotranslated texts in the case of Mansouri's works was not statistically significant. Regarding the second feature, the lexical density of authored texts was significantly higher than that of translated and pseudotranslated texts, but there was no significant difference between translated and pseudotranslated texts. Regarding the third feature of simplification, the statistical analyses showed that the sentence length of authored texts was significantly higher than that of pseudotranslated and translated texts, and the sentence length of pseudotranslated texts was significantly higher than that of translated texts.

Concerning the second research question, Laviosa's (2002) simplification hypothesis was not proved in terms of type/token ratio. However, regarding lexical density and mean sentence

length, the research findings verify the simplification hypothesis since translated texts were of lower lexical density and sentence length compared to authored texts. Regarding the simplification hypothesis' applicability to pseudotranslated texts, the research results suggest that the simplification hypothesis can be applied to pseudotranslated texts in terms of sentence length and lexical density since pseudotranslated texts, like translated texts, were of lower lexical density and shorter sentence length compared to authored texts. However, the simplification hypothesis did not apply to pseudotranslated texts in terms of type/token ratio as the results did not report any significant difference between pseudotranslated and authored texts. Overall, except for type/token ratio, the simplification hypothesis was applicable to the selected Persian translated and pseudotranslated texts.

6. Concluding Remarks

Overall, a comparative investigation of translated, authored, and pseudotranslated texts, a case study of Mansouri's works, revealed that the three groups of texts were of the same type/token ratio, i.e., the difference between their type/token ratio was not statistically significant. However, authored texts were of significantly higher lexical density than pseudotranslated and translated texts; but there was no significant difference between translated and pseudotranslated texts, which means they enjoy the same level of simplicity in terms of lexical density. With respect to sentence length, authored texts represented a higher sentence length than translated and pseudotranslated texts. Besides, the sentence length of pseudotranslated texts was significantly higher than that of translated texts. Therefore, Laviosa's (2002) simplification hypothesis was proved regarding lexical density and sentence length because translated texts were of lower lexical density and sentence length than authored texts. Similarly, the results approved the applicability of the simplification hypothesis to pseudotranslated texts in terms of lexical density and sentence length since pseudotranslated texts, like translated texts, were of lower lexical density and sentence length compared to authored texts, i.e., both translated and pseudotranslated texts were simpler than authored texts in terms of lexical density and sentence length. However, the results did not verify the simplicity of pseudotranslated and translated texts in terms of type/token ratio when compared to authored texts.

As noted earlier, pseudotranslators try to hide their works under the guise of translated texts by applying linguistic features of translated texts to their pseudotranslated texts (Shuttleworth & Cowie, 1997:134). Accordingly, it seems that Mansouri, as a distinguished author, successfully managed to apply such features to his own creative work. His success is evidenced by the fact that his works were widely received and sold as translations, with no significant challenge to his claim. This point suggests that the features he applied were sufficiently effective to be accepted as a translation by the general readership.

References

Anthony, Laurence (2018). AntConc (Version 3.5.7) [Computer Software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software/AntConc

Baker, Mona (1993). Corpus Linguistics and Translation Studies: Implications and Applications. In *Text and Technology: In honor of John Sinclair,* edited by Mona Baker, et al, 233-250. Philadelphia: John Benjamins.

Baker, Mona (1995). Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target, International Journal of Translation Studies 7(2): 223-243.*

Baker, Mona (1996). Corpus-Based Translation Studies: The challenges that Lie Ahead. In *Terminology, LSP and Translation: Studies in language engineering in honor of Juan C. Sager*, ed. by Harold Sommers, 175-186, Amsterdam: John Benjamins.

Baker, Mona and Gabriela Saldanha (2011). *Routledge Encyclopedia of Translation Studies* (second ed.). London and New York: Routledge.

Braheni, Reza (1985). *Kimia va khak* [Alchemy and Soil]. Tehran: Morghamin Publications.

Chafe, Wallace and Jane Danielewicz (1987). Properties of Spoken and Written Language. In M. Ghadessy (ed.), *Thematic Development in English Text,* 83-113, London: Pinter Publisher.

Browne, Edward. G (1893). *A Year Amongst the Persians: Impressions as to the Life, Character, & Thought of the People of Persia, Received During Twelve Months Residence in that Country in the Years 1887-1888*. London: Adams and Charles Black.

Chesterman, Andrew (2004a). Beyond the Particular. In *Translation Universals: Do They Exist?* Edited by A. Mauranen and P. Kujamäki, 33-49, Philadelphia: John Benjamins.

Chesterman, Andrew (2004b). Hypotheses about Translation Universals. In *Claim, Changes and Challenges in Translation Studies*, edited by G. Hansen, K. Malmkjaer and D. Gile, 1-14, Philadelphia: Benjamins Translation Library.

Chesterman, Andrew (2010). Translation Universals. In *Handbook of Translation Studies,* edited by Y. Gambier and L. Doorslaer, 175-180, Philadelphia: John Benjamins.

Even-Zohar, Itamar (1990). Polysystem Studies, *Special issue of Poetics Today*, 11:1.

Ghadessy, Mohsen and Yanjie Gao (2001). Simplification as a Universal Feature of the Language of Translation, *Journal of Asian Pacific Communication, 11(*1): 61-75.

Halliday, Michael, A. K. (1985/1994). *An Introduction to Functional Grammar*. London: Edward Arnold.

Johansson, Victoria (2009). Lexical Diversity and Lexical Density. In Speech and Writing: A Developmental Perspective, *Lund Working Papers in Linguistics, 53*: 61-79.

Kruger, Alet (2002). Corpus-based Translation Research: Its Development and Implications for General, Literary and Bible Translation, *Acta Theologica, 22* (1): 70-106.

Laviosa, Sara (2002). *Corpus-based Translation Studies: Theory, Findings, Applications*. New York: Rodopi.

Malmkjaer, Kirsten (ed.) (2004). *Translation in Undergraduate Degree Programs*. Philadelphia: John Benjamins.

Mansouri, Zabiollah (1984). *Islamic Sciences in the First Century*. Tehran: Chekameh Press.

Mansouri, Zabiollah (trans.) (1985). *The Lord of Alamoot*. Tehran: Donyaye danesh.

Munday, Jeremy (2012). *Introducing Translation Studies: Theories and Applications*. Routledge.

O'Sullivan, Carol (2010). Pseudotranslation. In Y. Gambier & L. Doorslaer, *Handbook of Translation Studies, 2:* 123-125, Philadelphia: John Benjamins.

Piccioni, Sara (2013). What Can Metaphor Tell Us about the Language of Translation, *Procedia- Social and Behavioral Sciences*, 95(3): 354-362.

Popovic, Anton (1976). *Dictionary for the Analysis of Literary Translation*. Alberta: University of Alberta Press.

Rambelli, Paolo (2009). Pseudotranslation. In M. Baker & G. Saldanha, *Routledge Encyclopedia of Translation Studies*, 208-211. Routledge.

Rizzi, Andrea (2008). When a Text Is Both a Pseudotranslation and a Translation: The Enlightening Case of Matteo Maria Boiardo (1441-1494). In A. Pym, M. Shlesinger and D. Simeoni (eds), *Beyond Descriptive Translation Studies: Investigations in Homage to Gideon Toury*, 153-161, John Benjamins publishing.

Salehyar, Gholam-hossein (1972). *Chehreh-ye Matbu'at-e Mo'aser* [The Face of Contemporary Journalism]. Press Agent.

Sharififar, Masoud, Zandrahimi, Mina and Fathimoghaddam, Fatemeh (2018). An Investigation of Bastin's adaptation as a kind of pseudotranslation: A case study of a year among Persians. *Iranian journal of English for Academic Purposes*, 7(1), 115-127.

Shuttleworth, Mark and Moira Cowie (1997). *Dictionary of Translation Studies*. Routledge.

Stubbs, Michael (1996). *Text and Corpus Analysis*. Oxford: Blackwell Publisher Ltd.

Toury, Gideon (1980). Translated Literature: System, Norm, Performance: Toward a TT-Oriented Approach to Literary Translation. In *In Search of a Theory of Translation, 35*-50, Tel Aviv: Tel Aviv University.

Toury, Gideon (1995). *Descriptive Translation Studies and Beyond* (vol.4). Philadelphia: John Benjamins.

Toury, Gideon (2012). *Descriptive Translation Studies and Beyond: Revised Edition* (Vol. 100). Philadelphia: John Benjamins Publishing.

Verma, Pritha and Gomma, A. S. Abdel-Salam (2019). *Testing Statistical Assumptions in Research*. New York: John Wiley & Sons.

Williams, Jenny and Andrew Chesterman (2002). *The Map: A Beginner's Guide to Doing Research in Translation Studies*. Manchester: St. Jerome Publishing.