

ARTICLE

Two translations of Mahfouz's *Awlad Haratina* (Children of our Alley): A computational-stylistic analysis

Mai Zaki¹, Emad Mohamed²¹ American University of Sharjah, ² Nazarbayev University

Keywords: literary translation, computational analysis, translator style, stylistics

<https://doi.org/10.22148/001c.116239>

Journal of Cultural Analytics

Vol. 9, Issue 3, 2024

Comparative studies of different translations for the same source text can be valuable sources of insights relevant to the fluid notion of 'translation style'. Such studies can employ a wide variety of techniques, including computational analysis which targets specific elements in the text in order to allow for a systematic view of translator style. This study attempts a computational-stylistic analysis of the two English translations of Naguib Mahfouz's controversial novel *Awlad Haratina* (literally, Children of our Alley). The aim of the study is two-fold. First, it aims to show how quantifiable computational and distant reading techniques can help identify patterns of stylistic differences between these two translations. Second, it attempts to situate the results of this analysis within the wider social context of the two English translations (Stewart 1981 and Theroux 1996) of one of the most famous modern Arabic novels. The results clearly show patterns of linguistic use specific to each of the two translations highlighting differences in lexical variety and richness, sentence structure, readability level, stylometric analysis as well some lexical choices. These results can be interpreted within the social context of producing those two translations, with particular reference to characteristics of retranslation as discussed in the literature.

1. Introduction

The novel under study by the Egyptian author and Nobel laureate Naguib Mahfouz is one of the most controversial novels in the history of modern Egyptian (and Arabic) literature. Originally published in serials in the Egyptian newspaper Al-Ahram in 1959, the highly allegorical novel which resonates with George Orwell's narrative style, caused ripples among the Egyptian readership and beyond with long-lasting effects. The novel ended up earning its author both condemnations from religious figures and praises from international institutions. It was specifically cited as one of the premises of awarding Mahfouz in 1988 the only Nobel Prize in Literature awarded to an Arab author until now. The Swedish Academy described it as "an unusual novel", in its press release announcing the prize (1988), where the main theme is "man's everlasting search for spiritual values". Depicting the religious and social history of man, the novel uses the narrative framework of a Cairo alley to chronicle developments from Genesis to modern times. With careful allegorical delineations of arguably God, Satan and a number of prophets,

the novel outraged religious leaders accusing the author of verging on heresy, and is often seen as a motivation for the assassination attempt on Mahfouz in 1994.

But the debates surrounding the novel since its publication in Arabic did not stop there; they also extended to its English translations. *Awlad Haratina*, which has been translated to many languages, is famously known to have two different English translations. The first one by Philip Stewart was published in 1981 by Heinemann under the name “Children of Gebelawi” Mahfouz. However, copyright issues exacerbated by the controversy surrounding its reception and motivated by its author’s post-Nobel Prize fame, led to another translation to be commissioned by the American University in Cairo. Thus, the second translation by Peter Theroux was published by Doubleday in 1996 under the title “Children of the Alley” Mahfouz.

Writers, journalists and researchers have been interested in discussing and chronicling the legacy of *Awlad Haratina* as well as the contexts of commissioning its two translations with a 15 years gap in between (Abu-Haidar; Ewais; Shoair). The two translations have also been the subject of several studies which analysed them from a wide variety of perspectives (Lundquist; Roded; Torabi and Seyed; El-Gabalawy; Hezam, among others). Moreover, and perhaps less commonly, the two translators themselves engaged in commenting on each other’s work in a somewhat lengthy and public debate (Stewart; Theroux, “Children of the Alley: A Translator’s Tale”; Theroux, “Naguib Mahfouz’s Children of the Alley”). However, very few attempted to actually highlight the stylistic differences between the two translations in a systematic way. Therefore, this study aims to fill this gap by applying stylometric and corpus-assisted distant reading techniques in order to compare the two translations and highlight stylistic similarities and differences between them. The quantitative and data-driven analysis sheds a new light on the English translations as texts and provides a different angle to viewing each translator’s style.

The paper tries to answer the following questions:

1. To what extent do the translations formally differ?
2. Can we use computational tools to tell the differences between the two translators and their respective styles?
3. What can the results tell us about re-translation and the general socio-cultural context of producing these two particular translations?

The paper is organised as follows: section 2 provides background information on the wider area of corpus studies of translation and translator style through a brief review of relevant literature. Section 3 briefly presents the two translations under study and the circumstances of their publication. Section

4 introduces the data and methodology, outlining the corpus-assisted and computational techniques utilised in the study. Section 5 presents the results of the analysis. Section 6 discusses the interpretation of the results as indicators of the stylistic identity of the translators as well as a reflection of the wider social context in which the two translations were produced. Finally, the paper ends with some concluding remarks on the affordances of the corpus-assisted distant reading approach in studying translator style.

2. Corpus studies and translator style

The huge developments in corpus linguistics in the last few decades meant that its intersecting points with other disciplines would become hot spots for research. The field of translation studies is no exception. In addition to corpus-assisted methodologies and corpus-based machine translation techniques which help in the activity of translation itself, corpus linguistics has also been increasingly beneficial to descriptive translation studies. There are several areas of interest in corpus-based translation studies, including research into translation universals (Baker, “Corpus Linguistics and Translation Studies: Implications and Applications”; Mauranen and Kujamäki), translation norms (Malmkjær, “Norms and Nature in Translation Studies”; Munday), language change (House; Kranich et al.) and individual variation among translators (Saldanha; Baker, “The Treatment of Variation in Corpus-Based Translation Studies”). This study follows the latter line of research strands.

The aim of corpus-based studies of individual variation in translations is to investigate translator style. The concept of translator style has been extensively discussed in relation to translational stylistics (Baker, among others). Generally speaking, translator style refers to the individual writing and translating choices made by a translator when translating a text from one language to another. This can include decisions about word choice, sentence structure, cultural references, idioms and the overall tone and style of the translation. According to Newmark, translator style is “the individual way a translator expresses the message of the SL text in the TL version, which may be determined by his register, his culture and his personality” (54). Venuti (*Translator's Invisibility*) expresses a similar opinion when he argues that the translator's style is a reflection of his or her ideological, cultural and professional identities.

Despite the previous definitions, or perhaps because of them, Saldanha acknowledges that the notion of style “is a notoriously vague concept” (26). Hence, she attempts to outline a methodological approach in defining it. According to her, there are mainly two approaches: (a) source text (ST)-oriented studies (e.g. Malmkjær; Boase-Beier); and (b) target text (TT)-oriented studies (e.g. Baker) of translator style. As the names suggest, studies in the first group focus on the style of the translation, rather than the translator, since it is defined by its relation to the ST. Yet, Boase-Beier (4)

BOX 1.1 THE ROLE OF STYLE IN TRANSLATION

- i. The style of the source text as an expression of its author's choices
- ii. The style of the source text in its effects on its readers (and on the translator as reader)
- iii. The style of the target text as an expression of choices made by its author (who is the translator)
- iv. The style of the target text in its effects on its readers

Figure 1. The role of style in translation according to Boase-Beier (89)

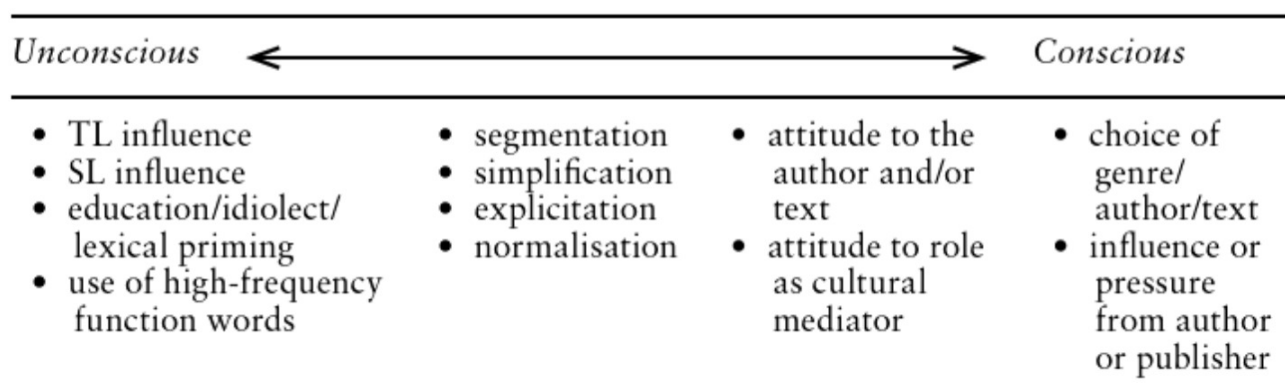


Figure 2. Continuum of elements of translator style according to Youdale (180)

also extends this view to outline four different ways style can have a role in translation (as Figure 1 illustrates) where it exploits the intersection of ST and TT on one hand, and author's/translator's choices and effect on readers on the other.

At the other end of this spectrum, TT-oriented studies of style focus entirely on the translated texts, usually by comparing features in two or more corpora of different translators without direct reference to the ST. Furthermore, Youdale (180) highlights some of the factors which can influence a translator's style seen as points on a continuum of conscious vs unconscious elements of a translator's style (see Figure 2). The nature of a continuum implies fuzzy distinction lines, and the author acknowledges this as he explains that "different aspects of style may be located at different points along the continuum, and these points themselves may vary from translator to translator and also vary over time" (Youdale 181).

Computer-assisted translation studies, i.e. those which use computational methods and tools to analyse the writing and translating choices made by translators, have transformed the study of translator style. Such studies

usually follow a specific design: a parallel corpus (monolingual or bilingual) and a set of “linguistic indicators”, as Zanettin (21) calls them, to act as formal operators in the study. Li et al. (106–08) and Youdale (183) provide some examples of corpus designs and formal operators which are also used in the domain of forensic stylistics. Some of the typical formal operators include:

- Type-token ration (TTR)
- Average sentence length (ASL)
- Lexical density
- Lexical/collocational creativity/richness
- Hapax legomena (words occurring only once)
- Distribution of most frequent words
- Analysis of specific patterns of choice (e.g. modal particles, report verbs, loan words, etc)
- Analysis of culture-specific items

An important caveat to note with this quantitative approach is that it does not aim to present an exhaustive description of a translator's style. As Li et al. explains, “the intent is not so much to find out what is the translator's style but rather whether she leaves thumbprints in the translated text, and if affirmative, what are some of the thumbprints” (109). This is what this study aims to do; i.e. to compare the translators' thumbprints in the two translations of this important literary work and to highlight some of the characteristic linguistic features for each translator. In doing so, we can capture what Baker calls the “individual profile of linguistic habits” for a translator.

3. Children of Gebelawi/the Alley: A background

Perhaps no other modern Arabic novel continues to conjure up debates since its initial publication in the late fifties until now as *Awlad Haratina*. El-Gabalawi (94) attempts an explanation as he argues that in this work “Mahfouz boldly explores areas of religious thought where most Egyptian writers fear to tread even in the contemporary age of relative liberalism.” However, while religious interpretations of the novel were primarily responsible for its blustery reception, others have recognised that it goes beyond that. (Hawthorne) writes in the New York Times, “religion is only one element of a larger critique of an entrenched patriarchal authoritarianism that extends from the familial to the political.” It is also worth mentioning that the publication and reception of the Arabic text was confounded from the start by the interplay of both religion and politics. One reflection of this is

that the publication of the original text serialised in the Egyptian newspaper Al-Ahram was only possible due to the interference of the Egyptian President at the time Gamal Abdel Nasser, while the first time the novel was published in a book was in 1962, not in its native country, but in Beirut, Lebanon. The novel's literal and allegorical layers of meaning associated with its characters and events are also engulfed in a deep sense of social protest, which can be deemed alarming or commendable depending on one's point of view. This combination of serious, even radical, ideas had a long-lasting impact on the novel and its interpretation, extending to its English translations as well.

As mentioned before, the first English translation of *Awlad Haratina* was published in 1981 under the name "Children of Gebelawi". The translator is Philip Stewart, the then young Oxford graduate who was commissioned to translate an already controversial work at the beginning of his career. Once commissioned with the translation, Stewart reportedly flew to Cairo and actually worked closely with Mahfouz himself during the process (but see Khalifa for a discussion of how both translators claimed access to the ST author). Another claim made by (Stewart), and backed up by (Allen), is that his translation is arguably the only complete version of the novel in any language, including Arabic, since the original Arabic manuscript was lost and some discrepancies have been noted both in the serialised version of Al-Ahram and the Beirut version.

After Mahfouz was awarded the Nobel Prize in literature in 1988, a relaunch of the English translation of the novel to the world stage was deemed necessary. Due to copyright complications, another translation was commissioned with a different translator. This time it was the American writer and translator Peter Theroux who undertook the task. His translation was published in 1996 under the title "Children of the Alley". A more experienced translator, Theroux has worked with other Arabic novels from different Arab countries including Iraq and Saudi Arabia. Theroux's translation would be technically considered a 'retranslation', a term often used in the literature to refer to any translation of a source text which follows a previous translation (by a different translator) in the same target language (Mathijssen; Van Poucke; Brownlie, among others). Interestingly, the translator himself, who noted that his translated book "has had a much more interesting life" than him (Theroux, "Children of the Alley: A Translator's Tale" 671), acknowledges that Mahfouz's "deeply spiritual and questioning novel [...] mirrors the stories of the Torah, Gospels and Koran"-an interpretation he shares with many of the novel's critics and one that Mahfouz himself never explicitly expressed.

Therefore, both translations of the novel appeared in testing times. Stewart's translation came out against the backdrop of a contentious reception of the original text in Egypt and the Arab world, while Theroux's translation came out in a rather complex socio-cultural context torn between international

praise and local condemnation. Several critical reviews/studies of both translations appear in the literature (e.g. Cachia; Ghorab; Khalifa, where some of them recognise the significance of the concept of retranslation in the discussion of this work. Although some studies (e.g. Susam-Sarajeva) argue that retranslation is a phenomenon which lacks a detailed or systematic study, there have been several attempts to define it. According to Baker ("Retranslation as Rewriting" 156), "retranslation is a practice that challenges the notion of a fixed original text and disrupts the hierarchies of linguistic and cultural dominance". Snell-Hornby (83) largely agrees with this characterisation and explains that "retranslation offers an opportunity to reflect on the status of the target language, its culture, and the implications of the previous translation". Indeed, much has been said about the relationship between a retranslation and the previous one, most notably regarding the motivations for producing a retranslation in the first place. Venuti (*Retranslations* 26), for example, sees a retranslation as largely triggered by competition, hence its need to assert its difference. He argues, in particular, that "retranslations are designed deliberately to form particular identities" (26) as he emphasises that retranslations are not just a form of rewriting, but of reinterpreting, the text.

While different studies have focussed on various aspects of the identities and assumptions associated with a retranslation, there has been a strong argument for the view that retranslations signal shifts in the ideological context of reception of a certain literary work. As Venuti explains, retranslations could be "a way of reshaping the target text to better suit the target culture" (Venuti, "Retranslations: The Creation of Value" 37). The case of the two English translations of *Awlad Haratina* would be a great example to test this hypothesis. Going beyond the traditional view of a necessary chronological ageing of a translation so that it requires a new one, Pym (82) highlights the "active rivalry" that defines multiple translations of the same source text. In his recent study, Khalifa (Khalifa, "The Hidden Violence of Retranslation: Mahfouz's *Awlād Ḥāratinā* in English" 147) takes this assumption even further and defines retranslation as "an act of defiance". Thus, he considers the case of the two translations of *Awlad Haratina* as "a site for struggle between competing translatorial agents." (145) He emphasises the rivalry between the two translators describing it as "combative interactions" (146) and elaborates on this through the analysis of, not the texts themselves, but the paratexts (introductions and commentaries by the translators) of the two translations. As mentioned before, the two translators did not shy away from commenting on each other's translations. While Theroux publicly described the translation of his predecessor as "perfectly adequate" and "pioneering" (Theroux, "Children of the Alley: A Translator's Tale" 668), he also referred to some "bad qualities" of previous Mahfouz translations (quoted in Khalifa).

This study brings back the focus to the actual texts of the two translations by employing computational and stylometric processes to highlight the distinct stylistic identities of the two translators. By interpreting these results, we also aim to shed light on some of the defining features of retranslations which ultimately qualify the relationship between the two translations of *Awlad Haratina*.

4. Data and methodology

To the best of our knowledge, there is currently no published research which specifically compares translator style in the two English translations of *Awlad Haratina* following a corpus-assisted or a computational approach. This study aims to rely on computational-stylistic analysis to inform our understanding and interpretation of the differences between the two translations. The analysis is largely data-driven as the aim is to neutralise all other variables and focus on translator style with the purpose, as Youdale (184) explains, “to perform an ‘x-ray’ of basic linguistic features which a close reading of even a short passage would be unlikely to reveal”. The analysis is based on two sub-corpora named: *Gebelawi* [Stewart] and *Alley* [Theroux] in the chronological order of their publication. After obtaining the digitised version of the two translations, some pre-processing was done to clean the data from digitising errors (fixing spacing errors, fixing mistakes in converting some letters, deleting running titles, etc). Then, the two translations were also aligned at chapter level to allow a more detailed comparison by Chapter. Therefore, the data was analysed in two forms: (a) as individual whole texts saved in separate sub- corpora; and (b) as comparable texts aligned at chapter level. Both translations follow the original division of the ST which includes 114 chapters spread over 5 parts.

The quantitative analysis includes several levels, which required the use of both corpus management software (Sketch Engine, Kilgarriff et al.) and machine learning techniques. In particular, the analysis explores the following aspects:

- Lexical variety (type-token ratio, lemma, hapax legomena)
- Syntactic / narrative style (sentence length and lexical bundles)
- Readability grade (level of comprehension difficulty)
- Stylometric analysis (bootstrap consensus trees, principal components analysis, multidimensional scaling and cluster analysis)
- Translator thumbprints (lexical choices and collocation)

5. The analysis

In this section, we present the results of the computational-stylistic analysis of the two English translations of Mahfouz's novel *Awlad Haratina*.

Table 1. Type-token ratios of the two English translations.

	Gebelawi	Alley	Difference ^a
Tokens*	145,207	147,980	+ 2773
Types	8,084	8,821	+737
Document TTR	0.05567	0.0596	+0.00393
STTR	0.4143	0.4145	+0.0002

*A token is defined as a whitespace delimited unit excluding punctuation (regular expression +)

^aAll “difference” in the tables refers to the quantitative difference of the second translation *Alley* (Theroux) compared to the first translation *Gebelawi* (Stewart).

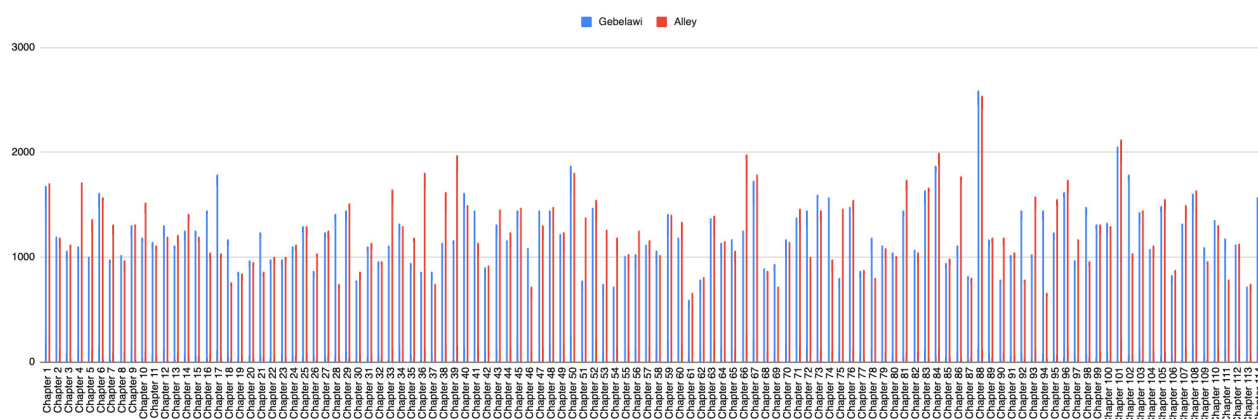


Figure 3. Comparing word count of the two translations by chapter

5.1 Lexical variety

The first step in the analysis is to look at word level. Table (1) presents the basic statistics of the type-token ratio (TTR) in the two corpora. TTR is a kind of vocabulary richness measurement which has been extensively used in authorship and genre analysis. This measurement is based on the simple (based on total words in the text) or standardised (based on every 1000 words in the text) ratio between the number of types (the unique word forms in a text) and tokens (the total words in a text). A higher TTR indicates that the author uses a wider range of vocabulary. To compute the TTR for the two translations, we used the standardised TTR where we divide each text into chunks of 1000 words each, compute the TTR for each chunk, then average the TTRs. This guarantees that the TTR is not affected by the size of the text since longer texts will result in lower TTR. The results are illustrated in Table (1).

The analysis was also expanded to compare text length variation by Chapter as illustrated in Figure (3). The results show that *Alley* has more word count in 73 chapters (64%). Table (2) below also shows the results of a detailed word count analysis by Chapter.

The results indicate that the American translator uses more vocabulary items than the British one. This may be due to the fact that the American translator was more experienced, and older, and it may also be due to having the

Table 2. Detailed word count analysis by Chapter

Text	Gebelawi	Alley	Difference
Chapters count	114	114	—
Mean	1,207	1,243	+ 36
STD	287	350	+ 63
25	1,006	1,001	- 5
50	1,170	1,190	+ 20
75	1,440	1,469	+ 29
Max	2,053	2,583	+ 530

Table 3. Hapax words in the two English translations

Text	Gebelawi	Alley	Difference
Hapax legomena	1,925	3,111	+ 1,186
As percentage of tokens	1.4	2.2	+ 0.8
As percentage of types	21.7	35	+ 13.3

advantage of the availability of the previous translation. One can also not rule out the fact that, by 1996, finding synonyms and more unique words was made easier by the availability of electronic methods.

Another kind of measurement for lexical variety is to identify the hapax legomena, i.e. words which occur in each corpus only once. Although this type of words could be overlooked due to their infrequency in the corpus, they are important for the analysis of lexical richness. As Table 3 illustrates, hapax words in *Alley* translation are more than *Gebelawi* by more than a thousand words.

However, to put this into perspective, it is useful to compare hapax words against the number of unique types. As Figure 4 shows, both translations have a little over 5000 words as common unique word types, but *Alley* corpus reflects a significantly higher percentage of hapax words to types indicating higher lexical richness.

5.2. Sentence length and lexical bundles

Moving to the sentence level, we look at two main features: sentence length and lexical bundles. We define a sentence as a string of words ending in a full stop, a question mark or an exclamation mark. Special care was taken of possible false friends resulting from dots in abbreviations. Both sentences and lexical bundles are important features which can reveal translation styles in corpus-assisted studies. Table 4 illustrates the results of comparing sentence length in the two translations (where sentences are words without punctuation). *Alley* has a little over 750 sentences more than *Gebelawi*, but the difference in standard deviation of the number of words per sentence is



Figure 4. Comparison of hapax words in unique types

Table 4. Sentence length of the two English translations

Text	Gebelawi	Alley	Difference
No. of sentences	10,465	11,244	+ 779
Mean (words per sentence)	13.8755	13.1608	- 0.7147
Median (words per sentence)	12	11	-1
Stnd. Dev.	8.7225	9.2074	+ 0.4849
Min.	1	1	—
Max.	77	85	+ 8

only +0.48 for *Alley*. Another indicator that *Alley* exhibits longer sentences is that the maximum number of words in a sentence is 77 in *Gebelawi*, but rises to 85 in *Alley*.

As for lexical bundles, or n-grams as they are also referred to in the literature, these were retrieved using the text analysis software Sketch Engine. The study focused on three-word up to six-word bundles which occurred a minimum of 5 times in either corpus (following Biber and Barbieri). Table 5 presents the results of the comparative analysis of the number of lexical bundles at each length.

The quantitative results in Table 5 reveal a few patterns. First, and consistent with other studies such as (Liu and Afzaal), the data reveals that the number of lexical bundles decreases as the bundle length increases. The big difference in frequency between three-word lexical bundles and all the others confirms the argument in (Biber and Barbieri) that three-word bundles are extremely common as collocational patterns compared to the longer ones. Second, the

Table 5. Lexical bundles in the two translations

No. of words	Gebelawi		Alley	
	Types	Tokens	Types	Tokens
Three	1,142	10,015	1,143	9,984
Four	184	1,426	178	1,313
Five	38	275	24	163
Six	6	51	1	5

Table 6. Example of lexical bundle expansion in *Gebelawi*

Length	Frequency						
3 n-gram	37			of	the	Great	
4 n-gram	37			of	the	Great	House
5 n-gram	7		wall	of	the	Great	House
6 n-gram	5	the	wall	of	the	Great	House

Table 7. Example of lexical bundle expansion in *Alley*

Length	Frequency						
3 n-gram	17	what	are	you			
4 n-gram	5	what	are	you	going		
5 n-gram	5	what	are	you	going	to	
6 n-gram	5	what	are	you	going	to	do

comparative results of the three-word bundles show that the number of types is almost the same in the two translations although the number of tokens in *Alley* (Theroux) is slightly lower. However, both the frequency and variety of the longer lexical bundles (four to six n-grams) are higher in *Gebelawi* (Stewart) compared to the other translation. This is especially notable in the six-word bundles which are rare in *Alley* translation. It is also worth mentioning that the shorter lexical bundles are often nested in the longer ones. Tables 6 and 7 illustrate how the shorter bundles are used as building blocks for the longer ones in both translations. Finally, this leads to the observation that the lexical bundles account for a slightly bigger proportion of the text in *Gebelawi* (8.5% of the whole text) than they do in *Alley* (8% of the whole text).

Lexical bundle frequency can be used as a proxy for language and translation creativity. An author who is more creative may be expected to use fewer fixed expressions. Translation poses further analysis challenges as the fewer lexical bundles may result from the differences between the SL and the TL, or from the translator's being creative.

5.3. Readability grades

Readability tests generally measure how easy a document can be read and understood. More specifically, such tests are designed to indicate comprehension difficulty for contemporary English by providing a numeric gauge of the comprehension difficulty level facing readers of a particular text. There are various types of readability tests, but they all have a common starting point: “to identify linguistic characteristics of texts that affect reading comprehension or, at least, are good predictors of it” (François 7). They result in a grade which reflects years of formal education necessary for a person to understand a particular text. Therefore, generally, a higher readability grade indicates a more challenging reading experience and vice versa.

In this study, nine different readability tests were applied as Table (8) illustrates. The results show that in five of these tests there is a significant difference in the score and grade level between the two translations, where the second translation *Alley* exhibits a higher, thus more difficult, readability grade. For example, the core measures in both the Flesch-Kincaid and the Gunning-Fog tests is word length, measured by the average number of syllables per word; and sentence length, i.e. the number of words per sentence (Solnyshkina et al.; Rodríguez Timaná et al.). In both of these tests the second translation *Alley* scored two points and three points higher respectively. The Dale-Chall test, on the other hand, relies on a list of 3000 familiar words which are considered easy to understand, while anything off that list will be counted as a difficult word (Chall and Dale). With this measure in mind, *Alley* also scored higher. The Automated readability index (ARI) depends on word length too as an indicator of difficulty, but it measures it in terms of a factor of characters per word instead of syllables as in other tests (De Heus and Hiemstra). Even using this method, the second translation scored two points higher. Finally, the Linsear Write test calculates difficulty based on a combination of three variables: (a) the average number of easy words (two syllables or less) per 100 words; (b) the average number of hard words (three syllables or more) per 100 words; and (c) the average number of sentences per 100 words (Wang et al.). In this test, *Alley* scored four points higher than *Gebelawi* on the readability grade.

5.4. Stylometric analysis

One aspect of the translators' style may be discovered through authorship (or translatorship) attribution. The main premise behind authorship studies is to use statistical and/or computational methods in order to measure some textual features which can then enable distinguishing between texts written by different authors (Stamatatos; El-Fiqi et al.). From a machine learning perspective, a stylometric analysis of a text depends on the processing of certain textual features with the purpose of quantifying the writing style of its author.

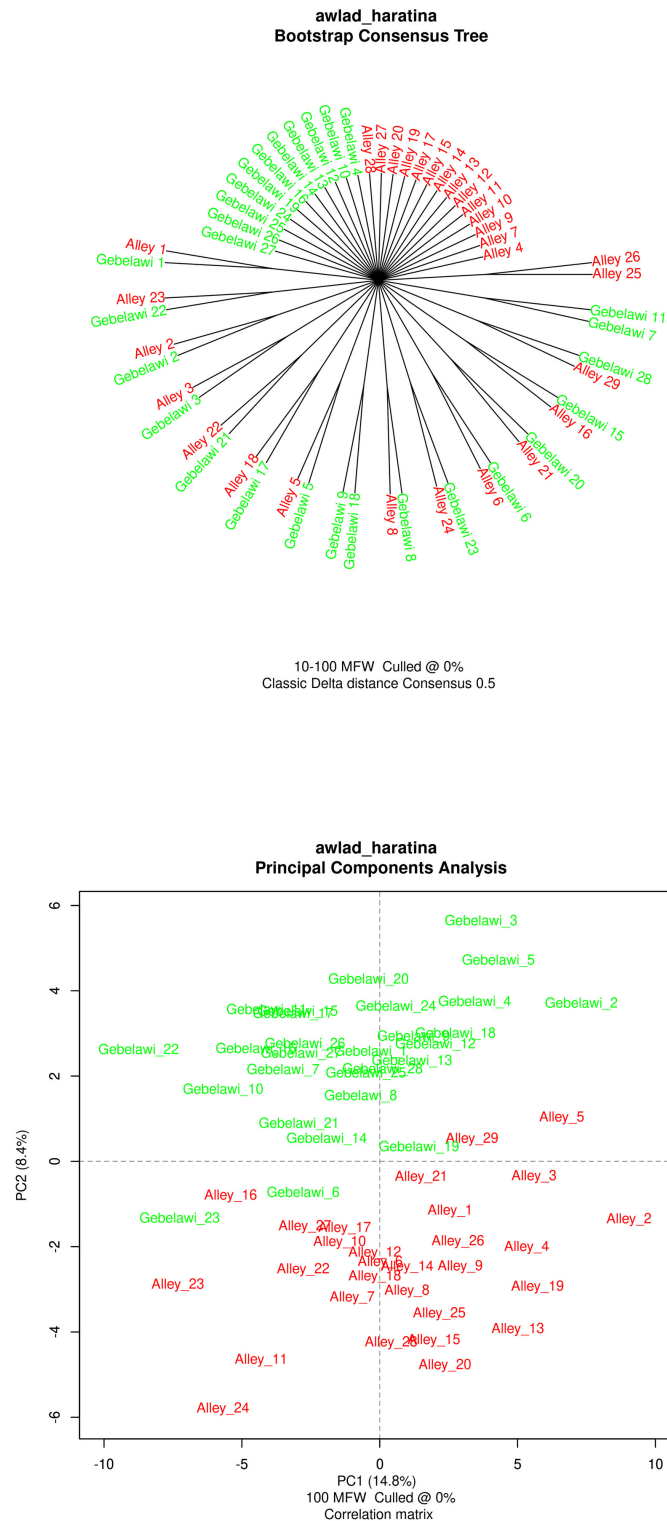
Table 8. Results of readability grade tests for both translations

Readability Tests	Gebelawi (Stewart)		Alley (Theroux)	
	Score	Grade-level	Score	Grade-level
Flesch_kincaid	6.2	6	8.23	8
Flesch	75.77	7	70.7	7
Gunning_fog	8.28	8	10.66	11
Coleman_liau	6.61	7	5.92	6
Dale_chall	7.37	9-10	8.54	11-12
ARI	5.51	6	7.3	8
Linsear_write	7.5	8	11.74	12
SMOG	11.34	11	11.2	11
Spache	5.26	5	6.26	6
Average score	14.87		15.62	

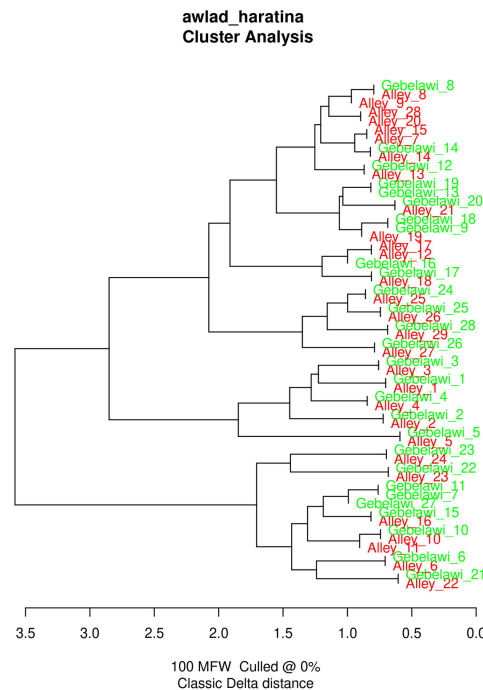
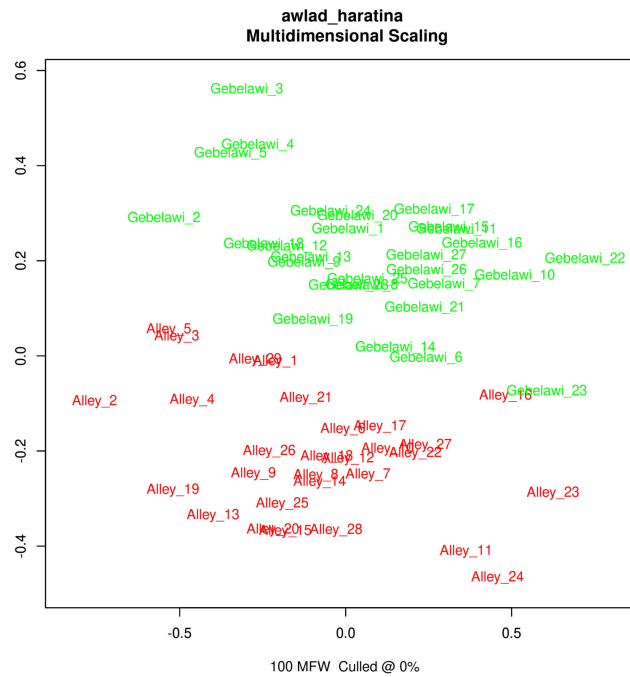
In this section, we use Stylo (Eder et al.), one of the main software packages for authorship attribution, to try and find out whether the two translators have distinct styles. Within authorship studies, style is generally understood as the distribution of most frequent words. Stylo algorithms depend on word or character n-grams and processes input data several times in order to establish links between texts. We have run experiments using Bootstrap Consensus Trees, Principal Components Analysis (PCA), Multidimensional scaling and Cluster Analysis and the resulting visualisations reflect the relative distance between the two texts in a 2-dimensional space.

The results from the cluster analysis and bootstrap consensus trees are not conclusive, i.e. some chapters from both translations seem to be very similar while others are more distinct. However, the results from both PCA and multidimensional scaling show a clearer distinction between the two translations which reflects a stylistic difference that can be computationally distinguished as illustrated in Figures (5)-(8).

We get similar results when we use supervised machine learning to predict whether a specific piece of text was the output of a specific translator. In these experiments, we use the top 100 words as features. To make things more consistent, we have normalised the characters' names so these may not be clues for the machine learning algorithms. For example, all forms of "Gebel" were converted to "Gabal". We then extracted the most frequent 100 words overall and used as features the frequencies of these top words in both translations. Each translation was divided into sections of 500 words and the feature values are the frequency of each of the top 100 words in each of these two translations. The target was then whether the vector belonged to *Alley* or *Gebelawi*. We used the Extreme Gradient Boosting (XGBoost) classification algorithm in a 10-fold cross validation setting and we used the F1 score for evaluation. We obtained an average F1 score of 0.775, which means that supervised machine learning can, in most cases, tell the difference between these two translations of the novel. This also indicates that each of



the translators has his own style. However, although the stylometric methods used here are useful, we acknowledge that they have several limitations. For example, we have only relied in the analyses either on function words or the top 100 words in each section, but we have not utilised other features such as discourse markers or syntactic differences (which needs an additional layer of dependency parsing in the analysis).



5.5. Translator thumbprints

The linguistic thumbprint of a translator can be further characterised through some of their vocabulary choices, distribution and collocations. For example, the analysis looks into the frequency and distribution of the names of the main characters in both translations. The novel is divided into five parts, based on the main characters in the plot. The first four have been

Table 9. Frequencies of the names of main characters in both translations

Name	Gebelawi	Alley	Difference
Gebelaawi / Gabalawi	222	275	+ 53
Adham	341	354	+ 13
Gebel / Gabal	494	527	+ 33
Rifaa	343	401	+ 58
Qaasim / Qassem	426	425	- 1
Arafa	279	265	- 14

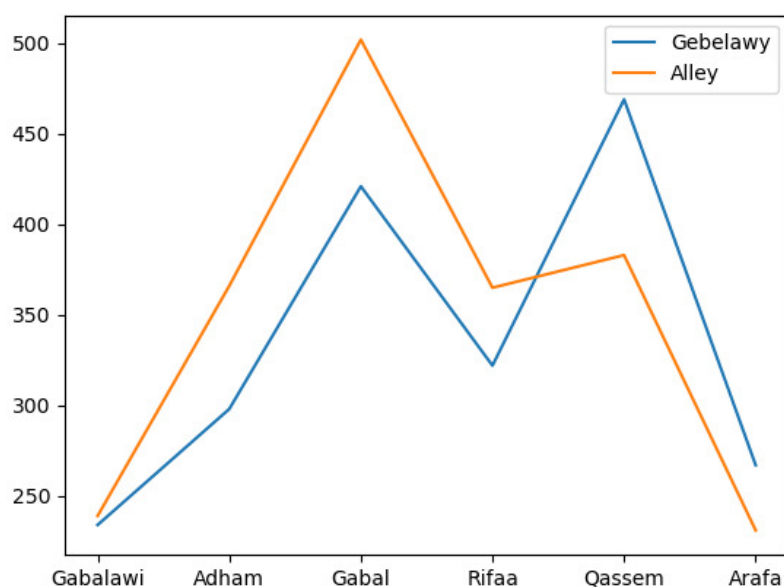


Figure 9. Comparing frequencies of character names in both translations

associated with prophets as they illustrate in their own stories the progression of mankind on earth starting with Adham (Adam), going through the three messengers of Judaism, Christianity and Islam: Gabal (Moses), Rifaa (Jesus), Qassem (Muhammad). Then the last character represents a symbol of change, i.e. Arafa (modernisation). The other important starting figure in the novel is Gebelawi, who has been interpreted as a symbol of God. A corpus search of these names in the two translations as they are spelled results in the frequencies illustrated in Table 9. It shows a pattern for the second translation to refer to names of characters more frequently.

A more comprehensive look at these character names in the two translations normalises all spelling variations (e.g. Gebelaawi/Gabalawi, Qaasim/Qassem, etc) and includes all grammatical variations (e.g. Adham/Adham's). The results, illustrated in Figure (9), still show a pattern in the second translation *Alley* to use character names more frequently with the exception of the two characters Qassem and Arafa.

Upon close examination of the cases where the second translation refers to character names more frequently, it turns out that there are two main reasons why this happens. First, in some cases, the character name appears in *Alley* whereas *Gebelawi* uses a pronoun (he/him) instead. Secondly, which is the more common case, *Gebelawi* does not mention the character's name because the whole phrase is deleted in the translation, whereas *Alley* follows the ST more closely. In examples (1) and (2) below, the character names Rifaa and Gabal only appear in the *Alley* translation.

EXAMPLE (1):

فقال معتذراً: لدي عمر كامل للنجارة، ولكن يهمني الآن أن أزور المقاهي جميعاً“

“– I have a whole lifetime for carpentry, but just now I’m anxious to visit all the cafes.” (*Gebelawi*, Ch. 46)

“I have a whole life for carpentry,” *Rifaa said in a tone of apology*, “but for now I want to visit all the coffeehouses.” (*Alley*, Ch. 46)

EXAMPLE (2)

وخرج جبل من صمته قائلاً: لعلها جريمة مزعومة لم تقع؟ واندفع الغضب في صدر زقلط لدى سماعه صوت

جبل فقال: لا ينبغي أن نضيع الوقت في الكلام.

"Gebel broke his silence: — Perhaps it’s only an allegation.

Thudclub was furious at this, and he said: – There’s no need for us to waste time talking.” (*Gebelawi*, Ch. 31)

“Maybe it’s just a story—maybe this crime didn’t take place!” *Gabal* said, breaking his silence.

Rage ignited in Zaqlut’s breast at the sound of *Gabal’s* voice. “We must not waste our time with talk.” (*Alley*, Ch. 31)

However, in example (3), it is noted that even though the name Gabalawi appears in *Alley* but not in *Gebelawi*, it is not explicitly mentioned in the ST either but was a result of an enriched interpretation on part of the translator.

EXAMPLE (3)

“ونال نصيبه كاملاً من ريع الوقف ورأى الأبنية تشيد باسم الوقف ثم تتوقف بأمر قدري”

“..received his share of the Trust’s revenue and watched the new buildings go up, and then to have seen all this stop by order of Qadri.” (*Gebelawi*, Ch. 97)

“He got his full share of the estate revenue, and saw the buildings built in *Gabalawi’s* name, and then stopped on Qadri’s orders.” (*Alley*, Ch. 97)

In addition to character names, the analysis looked into the collocation of some high frequency words in the two translations¹. The events in the novel all revolve around the people living in the “alley”, i.e. a symbolic reference to earth, and a word which also appears in the title of the second translation. The word “alley” is the second most frequently used noun in both translations, with Theroux's *Alley* using it 117 times more than *Gebelawi* translation (*Alley* = 668 times / *Gebelawi* = 551 times). The examples below illustrate the differences in the use of this word in the two translations:

EXAMPLE (4)

“وإذا بأصوات نساء تتطلق من النوافذ في ضوضاء غير متميزة ضاعت في ضوضاء المعركة، غير أن النساء بدون وهن يشرن بأيديهن في فزع تارة نحو طرف الحارة الشرقي وتارة نحو الطرف الآخر.”

“Women began shouting from the windows, but their cries were lost in the noise of the battle. However, they could be seen pointing in horror, now to the east and now to the west.” (*Gebelawi*, Ch. 90)

“Women's voices now shrieked from windows, a noise that could not be heard above the chaos of the battle, though they could be seen pointing in terror, now to the east end of the **alley**, now to the other end.” (*Alley*, Ch. 90)

EXAMPLE (5)

“لماذا عدت بنا إلى هذه الحارة؟ فقال عرفة والابتسامة مازالت في شفثيه: في كل مكان أسمع هذا الكلام وهذه حارتنا على أي حال، وهي الحارة الوحيدة التي يمكننا الإقامة بها..”

“Why did you bring us back to this *place*? Arafa went on smiling. I hear this sort of talk everywhere, and anyway this is our *Alley*. It's the only *one* we can live in..” (*Gebelawi*, Ch. 93)

“Why did you bring us back to this *alley*?” his companion whispered irritably in his ear.

Arafa smiled still as he answered. “I hear this kind of talk everywhere, and anyway this is our *alley*, and this is the only *alley* we can live in.” (*Alley*, Ch. 93)

EXAMPLE (6)

“وذهب حنش إلى مستوقد الصالحية وسأل عن زبال حارة الجبلوي، ثم سأل عن زباله الحارة

“Hanash went to the furnace and found the garbage man of Gebelawi **Alley**. He asked him about the rubbish.” (*Gebelawi*, Ch. 114)

¹ See Appendix for lists comparing the top 10 most frequent content words in the two translations.

“And so Hanash went to the dump at Salihiya and asked for the Gabalawi **Alley** garbageman and asked him about the trash from the **alley**.” (*Alley*, Ch. 114)

In those examples, we see the second translation *Alley* following the ST more closely in its depiction of the “alley” as a central place in the narrative represented by the frequent reference to it, unlike Stewart’s translation who was happy to omit or replace the noun “alley” in many instances. But the differences do not stop there. Figures (10) and (11) below illustrate the various collocations of this noun in the two translations:

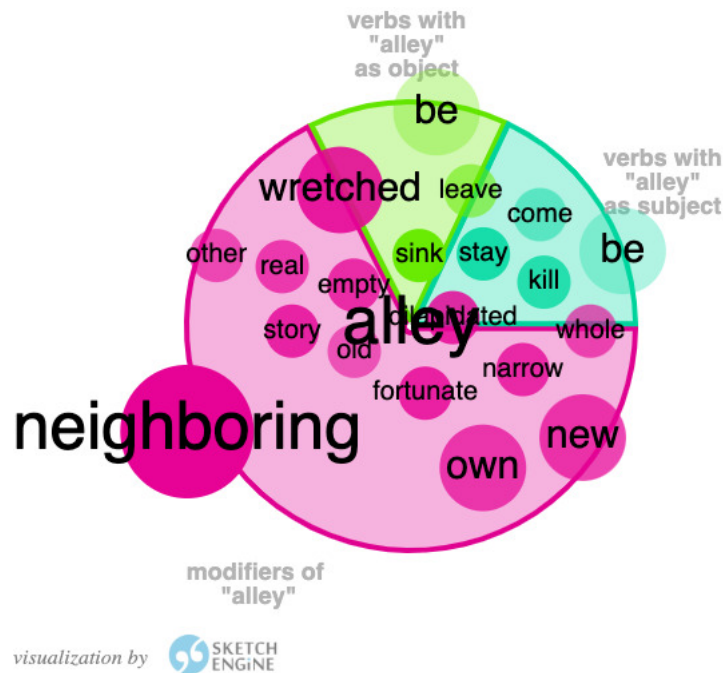
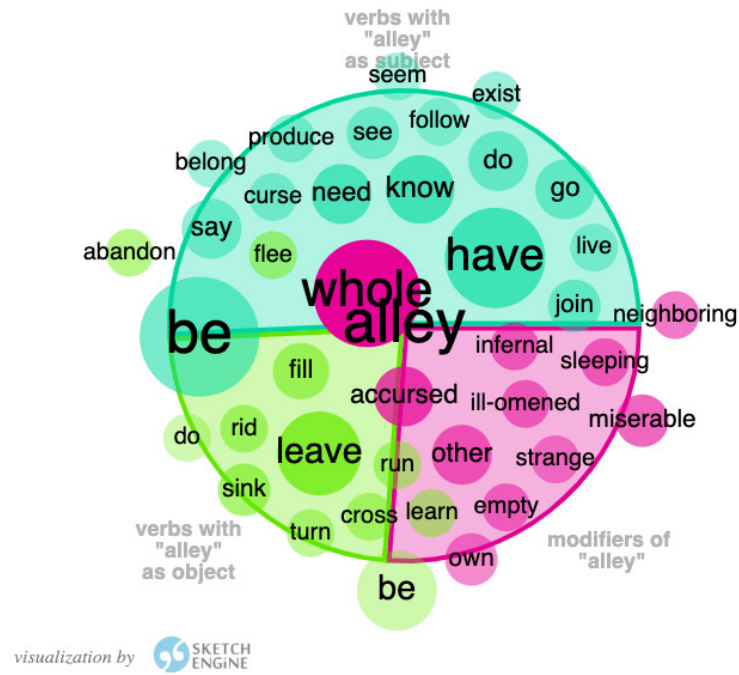


Figure 10. Collocations of the noun “alley” in *Gebelawi*

The most noticeable feature is that in *Gebelawi*, the highest frequency collocations of “alley” are in the form of modifiers which constitute almost two thirds of these collocates. This is contrasted with the collocates in *Alley*, where half of them are verbs for which the noun “alley” is a subject. So, not only is the concept of “alley” more present in the second translation, but it is also associated with more agency in the form of the number of high frequency verbs which collocate with it. This is also highlighted in the frequencies of other relevant n-gram structures such as “our alley”, “of the alley” and “in the alley” as the chart below illustrates:

Figure 11. Collocations of the noun “alley” in *Alley*

Gebelawi and Alley

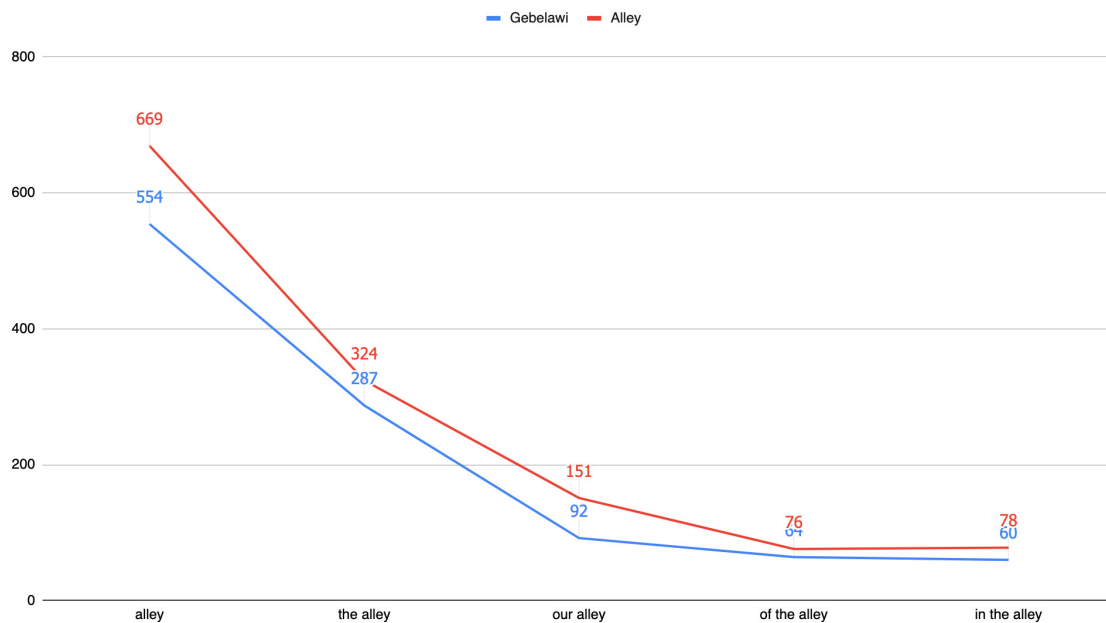


Figure 12. Frequency of “alley” and related structures in the two translations

6. Summary and discussion

The analysis in this study has computed several linguistic features at various levels in the two translations in order to tease out their unique characteristics. The main purpose is to illustrate how quantifiable features such as types, tokens, hapax words, sentence length, lexical bundles, readability grades, collocations, etc. can act as discovery tools for identifying patterns of translator style. The results clearly highlight some stylistic divergences between the two translations. This is a summary of the results:

- *Alley* is a longer translation as it has more overall word count and the word count by Chapter is higher in 64% of Chapters than *Gebelawi*.
- *Alley* reflects higher lexical richness as it has 1000+ hapax words more than *Gebelawi*, and the percentage of hapax words to overall types (35%) is also higher than *Gebelawi* (21.7%)
- *Alley* has higher count and longer sentences, with the maximum number of words in a sentence in *Alley* being 85, compared to 77 in *Gebelawi*.
- *Gebelawi* has higher frequency and variety of the longer lexical bundles (4 to 6 n-grams), and lexical bundles account for a higher proportion of the overall text in *Gebelawi* compared to *Alley*.
- *Alley* has higher scores in 5 out of 9 readability tests (Flesch_kincaid, Gunning_fog, Dale_chall, ARI and Linsear_write) indicating a more difficult reading experience.
- Two of the stylometric analyses performed (PCA and multidimensional scaling) clearly showed that both translations are stylistically distinct, while two other computational experiments (cluster analysis and bootstrap consensus tree) were inconclusive.
- Supervised machine learning techniques using top 100 words as features can distinguish between the two translations.
- *Alley* refers to character names more frequently in line with the ST, while *Gebelawi* uses more pronouns or deletion of character names.
- *Alley* has higher frequency of the noun “alley” and its relevant n-grams (“the alley”, “our alley”, “in the alley”, “of the alley”) compared to *Gebelawi*.
- *Alley* uses the noun “alley” as a subject with more verb collocates, while *Gebelawi* uses the noun “alley” with more modifiers.

In order to interpret these results, we will have to refer to two factors working in parallel: the concept of retranslation and the social context of producing the two translations. In terms of lexical variety and richness, these results seem to be consistent with assumptions regarding retranslations which are typically produced in direct competition with an earlier translation. In this case, given the copyright problems which engulfed the first translation and motivated the need for a second one, it could be argued that the second translation of *Awlad Haratina* was meant to be ‘better’, i.e. lexically richer, longer and more stylistically appealing than the first. This is supported by the analysis of word count, hapax words and sentence length which shows that *Alley* has higher numbers in all of these categories.

Additionally, *Alley* also scored higher in 5 out of the 9 readability tests applied in the study. These results reflect how the variables of vocabulary and sentence structure can affect the readers' experience while reading a particular text. Although these tests employ different mathematical formulae in measuring difficulty, they mostly rely on quantifiable features such as word length, sentence length and frequency. Therefore, the results are not only another indicator of differences in translators' style, but they also tie in with one of the early hypotheses on retranslations, i.e. that first translations, compared to subsequent ones, are generally motivated by creating an easy level of readability (Gambier 414). Therefore, the lower scores on the readability tests for *Gebelawi* reflect a simpler, more manageable reading experience, which is supported by a shorter text with lower lexical variety and more frequent lexical bundles. Indeed, the translator himself implies the intention to create an easy read as he says in the introduction to his translation that it "is aimed at the general reader with no prior knowledge of the Arab world; no words have been used that cannot be found in a good English dictionary" (Stewart 59).

The only time *Gebelawi* scored higher was in the lexical bundles analysis. Lexical bundles have been described as "building blocks" in discourse (Biber and Barbieri 277) which have a statistical tendency to co-occur but are not easily detectable without the aid of computers. Analysis of lexical bundles in translation studies have been growing since Baker's (Baker, "The Treatment of Variation in Corpus-Based Translation Studies") pioneering work, and they have been considered a useful measure of linguistic differences between translations of the same source text. According to Lee (382), translated texts are expected to rely heavily on high-frequency lexical bundles "if they are geared toward fluency and readability". It seems that *Gebelawi* achieved this as the analysis shows Stewart's translation was constructed as an easier text to read. However, with the attention the Arabic novel gained in the post-Nobel era, this perhaps put more pressure on the second translation to be different. The focus seems to have shifted from creating an easy read to creating a translation that is closer to the ST and one that is more stylistically nuanced. In this sense, Theroux's translation seems to fit in with the assumption that retranslations are expected to be more attentive to the style of the ST² (Gürçağlar), as was also supported by the analysis of character names in the two translations.

The distant reading approach, enabled by corpus analysis tools, was used in this study to help in painting a clearer picture of the translators' style by tracing some of their linguistic thumbprints. We looked into the use of character names and some high frequency words/structures to reveal

2 Although see Greenberg (202) who comments that Theroux's translation "while a major accomplishment in translation, cannot help but obscure some of the author's more subtle tricks" in reference to the translation of certain culturally-complex lexical items such as *mu'allim* (master) or *futuwwa* (strongman).

a few differences between the translations. The analysis of these samples from the text seem to show that Theroux's translation follows the ST more closely both in its depiction of characters and the centrality of the "alley" in the narrative space as reflected in the frequency and distribution of these linguistic items. The differences in the types of collocations of the word "alley" in the two translations also shows that the concept of "the alley" has more agency and dynamism in the second translation being the subject of numerous verbs, while it collocates with more modifiers in the first translation indicating a rather descriptive tone. This argument could be perhaps extended to explain the change in the title itself from *Children of Gebelawi* in the first translation to *Children of the Alley* in the second to be seen as more than just a necessary differentiating feature of a re-translation. That is, it could be perceived as projecting a focus of the narrative onto the place as well as its people. While further analysis needs to be done with regards to this point, these initial results are another indication of a stylistically distinctive choice on part of the translators and can pave the way for literary critics to support/refute their arguments.

Finally, the stylometric analysis of the two translations, which is based on the distribution of the most frequent words in the texts (at chapter level), showed that the two can be computationally distinguished in two of the tests employed. Therefore, machine learning techniques were shown to successfully distinguish between the two translations, and illustrated that textual metrics are strong indicators of translator style in line with previous studies (e.g. Lynch and Vogel, Claesen). Pooling all these quantitative results, this study aims to show that computational and data-driven studies of translated texts can complement descriptive studies which are grounded in literary or translation theories and need to be interpreted within the larger social and cultural contexts of the production and reception of the translation. In this sense, it becomes increasingly important to "understand the specific constraints, pressures and motivations that influence the act of translating and underlie its unique language" as Laviosa (478) points out. That is why the translation and retranslation of *Awlad Haratina* can be the perfect example for the study of such constraints and pressures for which no single factor can assume sole responsibility. This is a case of a novel laden with claims of religious or political allegorical interpretations and with two translations divided by a world-class prize that brought the ST to global centre stage. Therefore, the language in both translations would surely be influenced by multiple factors, which Pym (146) described as "the entourage of patrons, publishers, readers and politics". The quantitative study of these two translations thus complements (Khalifa)'s focus on the sociology of translation which includes an examination of the various agents involved including the publishers, the translators and the paratext of both translations.

7. Concluding remarks

Decades after the first publication of *Awlad Haratina* and after being translated into many languages, it still proves to be a work of many dimensions, and definitely one which defies one single interpretation. According to Venuti (34), “translations are profoundly linked to their historical moment”. This is truly the case with these two translations, which, as many critics noted, are different because they were commissioned and published in very different, yet equally stressful, circumstances thanks to the controversy engulfing the source text. This study presented a corpus-driven stylistic analysis of the two English translations of Mahfouz's *Awlad Haratina* with the aim to show that quantifying stylistic features in a systematic way can shed some light on translator's styles and contribute to the study of retranslation in the context of this particular novel.

The very notion of translator style has proved to be an elusive phenomenon. However, computational and corpus-driven studies have helped systematise the study of style by isolating those linguistic features which can reveal the translator's thumbprints or, as Youdale (178) calls them, the “unconscious translation habits”. Such habitual patterns of linguistic usage, examined through variables such as frequency, word and sentence length, lexical bundles, etc., provide a solid starting point for answering questions related to both the production and reception of any translation, and especially in the cases of comparing translations of the same text. Even when dealing with such a giant figure in modern Arabic literature like Mahfouz and arguably his most famous novel, this study wanted to show that a translation can never simply be a vehicle for the reproduction of the authorial style of the ST. By focusing on the texts of the two translations and teasing out the specific patterns in their respective texts, the study hopes to contribute to the area of computational translation stylistics in general and to studies of *Awlad Haratina* in particular as a unique literary work both in its original language and its translations.

Submitted: November 13, 2023 EDT, Accepted: December 13, 2023 EDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

WORKS CITED

- Abu-Haidar, Jareer. "Awlad Haratina' by Najib Mahfuz: An Event in the Arab World." *Journal of Arabic Literature*, vol. 16, no. 1, 1985, pp. 119–31. Crossref, <https://doi.org/10.1163/157006485x00112>.
- Allen, Roger. "The Happy Traitor: Tales of Translation." *Comparative Literature Studies*, vol. 47, no. 4, Jan. 2010, pp. 472–86, <https://doi.org/10.5325/complitstudies.47.4.0472>.
- Baker, Mona. "Corpus Linguistics and Translation Studies: Implications and Applications." *Text and Technology: In Honor of John Sinclair*, edited by M. Baker et al., 1993, pp. 233–50.
- . "Retranslation as Rewriting." *Routledge Encyclopaedia of Translation Studies*, edited by M. Baker and G. Saldanha, Routledge, 2011, pp. 156–59.
- . "The Treatment of Variation in Corpus-Based Translation Studies." *Language Matters*, vol. 35, no. 1, Jan. 2004, pp. 28–38. Crossref, <https://doi.org/10.1080/10228190408566202>.
- . "Towards a Methodology for Investigating the Style of a Literary Translator." *Target. International Journal of Translation Studies*, vol. 12, no. 2, Dec. 2000, pp. 241–66. Crossref, <https://doi.org/10.1075/target.12.2.04bak>.
- Biber, Douglas, and Federica Barbieri. "Lexical Bundles in University Spoken and Written Registers." *English for Specific Purposes*, vol. 26, no. 3, Jan. 2007, pp. 263–86, <https://doi.org/10.1016/j.esp.2006.08.003>.
- Biber, Douglas, et al. *Longman Grammar of Spoken and Written English*. Longman London, 1999.
- Boase-Beier, Jean. *Stylistic Approaches to Translation*. Routledge, 2014, <https://doi.org/10.4324/9781315759456>.
- Brownlie, Siobhan. "Narrative Theory and Retranslation Theory." *Across Languages and Cultures*, vol. 7, no. 2, Dec. 2006, pp. 145–70. Crossref, <https://doi.org/10.1556/acr.7.2006.2.1>.
- Cachia, P. "Naguib Mahfouz: Children of Gebelawi (Book Review)." *Journal of Arabic Literature*, vol. 14, no. 1, 1983, pp. 101–02, <https://doi.org/10.1163/157006483x00123>.
- Chall, Jeanne S., and Edgar Dale. *Readability Revisited: The New Dale-Chall Readability Formula*. 1995.
- Chesterman, Andrew. *Memes of Translation: The Spread of Ideas in Translation Theory*. John Benjamins, 2004.
- Claesen, Thomas. *Style and Coordination in Literary Translation: A Case Study of the Coordination Style in Joseph Roth's Hotel Savoy and Its Two Dutch Translations*. Ghent University Dissertation, 2014.
- Cronin, Michael. *Translation and Globalization*. Routledge, 2003.
- De Heus, Marije, and Djoerd Hiemstra. "Readability of the Web: A Study on 1 Billion Web Pages." *CEUR Workshop Proceedings*, vol. 986, 2013.
- Eder, Maciej, et al. "Stylometry with R: A Package for Computational Text Analysis." *The R Journal*, vol. 8, no. 1, 2016, p. 107, <https://doi.org/10.32614/rj-2016-007>.
- El-Fiqi, Heba, et al. "A Computational Linguistic Approach for the Identification of Translator Stylometry Using Arabic-English Text." *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, June 2011, pp. 2039–45, <https://doi.org/10.1109/fuzzy.2011.6007535>.
- El-Gabalawy, Saad. "The Allegorical Significance of Naguib Mahfouz's Children of Our Alley." *International Fiction Review*, vol. 16, no. 2, 1989, pp. 91–97.
- Ewais, K. "Najīb Mahfūz Rafada I'ādat Tib'āt Awlād Hāratinā Khawfan 'alā Usratihi." *Al-Arabiya*, 2006, <http://www.alarabiya.net/articles/2006/08/30/27047.html>.

- François, Thierry. "When Readability Meets Computational Linguistics: A New Paradigm in Readability." *Revue Française de Linguistique Appliquée*, vol. 20, no. 2, Nov. 2015, pp. 79–97, <https://doi.org/10.3917/rfla.202.0079>.
- Gambier, Yves. "La retraduction, retour et détournement." *Meta*, vol. 39, no. 3, 1994, p. 413, <https://doi.org/10.7202/002799ar>.
- Ghorab, Muhammad. *One Text, Two Receptions: Mahfouz's Children of the Alley*. 2013. American University of Sharjah, MA thesis.
- Greenberg, Nathaniel. "Naguib Mahfouz's Children of the Alley and the Coming Revolution." *The Comparatist*, vol. 37, no. 1, May 2013, pp. 200–18, <https://doi.org/10.1353/com.2013.0020>.
- Gürçaglar, Şehnaz Tahir. "Retranslation." *Routledge Encyclopedia of Translation Studies*, edited by Mona Baker and Gabriela Saldanha, 3rd ed., Routledge, 2009, pp. 484–90, <https://doi.org/10.4324/9781315678627-103>.
- Hawthorne, Melanie. "Seasons of the Patriarch." *New York Times*, 1996, <https://www.nytimes.com/1996/02/18/books/seasons-of-the-patriarch.html>.
- Hezam, Abdulrahman Mokbel. "Children of the Alley: Mahafouz's Allegory of Violence and Oppression." *Arab World English Journal (AWEJ)*, no. 3, 2016, pp. 86–94, <https://doi.org/10.2139/ssrn.2843960>.
- House, Juliane. "Beyond Intervention: Universals in Translation." *Trans-Kom*, vol. 1, no. 1, 2008, pp. 6–19.
- Khalifa, Abdel-Wahab. "The Hidden Violence of Retranslation: Mahfouz's *Awlād Ḥāratinā* in English." *The Translator*, vol. 26, no. 2, Apr. 2020, pp. 145–62, <https://doi.org/10.1080/13556509.2020.1751472>.
- . *The Socio-Cultural Determinants of Translating Modern Arabic Fiction into English: The (Re)Translations of Naguib Mahfouz's Awlad Haratina*. 2017. University of Leeds, Unpublished Doctoral Dissertation.
- Kilgarrieff, Adam, et al. "The Sketch Engine: Ten Years On." *Lexicography*, vol. 1, no. 1, July 2004, pp. 7–36. *Crossref*, <https://doi.org/10.1007/s40607-014-0009-9>.
- Kranich, Svenja, et al. "A Tentative Typology of Translation-Induced Language Change." *Multilingual Discourse Production: Diachronic and Synchronic Perspectives*, edited by Svenja Kranich, John Benjamins Publishing Company, 2011, pp. 11–44, <https://doi.org/10.1075/hsm.12.02kra>.
- Laviosa, Sara. "The Corpus-based Approach: A New Paradigm in Translation Studies." *Meta*, vol. 43, no. 4, 1998, p. 474, <https://doi.org/10.7202/003424ar>.
- Lee, Christopher. "Using Lexical Bundle Analysis as Discovery Tool for Corpus-Based Translation Research." *Perspectives*, vol. 21, no. 3, Sept. 2013, pp. 378–95, <https://doi.org/10.1080/0907676x.2012.657655>.
- Li, Defeng, et al. "Translator Style: A Corpus-Assisted Approach." *Corpus Methodologies Explained*, edited by Minghua Ji et al., Routledge, 2016, pp. 113–46.
- Liu, Kanglong, and Muhammad Afzaal. "Translator's Style Through Lexical Bundles: A Corpus-Driven Analysis of Two English Translations of Hongloumeng." *Frontiers in Psychology*, vol. 12, May 2021, pp. 1–10, <https://doi.org/10.3389/fpsyg.2021.633422>.
- Lundquist, Susanne. "Narrative Theory in Naguib Mahfouz's The Children of Gebelawi." *Understanding Others: Cultural and Cross-Cultural Studies and the Teaching of Literature*, edited by Joseph Trimmer and Tilly Warnock, 1992, pp. 213–29.

- Lynch, Gerard, and Carl Vogel. "The Translator's Visibility: Detecting Translational Fingerprints in Contemporaneous Parallel Translations." *Computer Speech & Language*, vol. 52, Nov. 2018, pp. 79–104. Crossref, <https://doi.org/10.1016/j.csl.2018.05.002>.
- Mahfouz, Naguib. *Awlad Haratina*. Dar Al-Adab, 1967.
- . *Children of Gebelawi*. Translated by Philip Stewart, Heinemann, 1981.
- . *Children of the Alley*. Translated by Paul Theroux, Doubleday, 1996.
- Malmkjær, Kirsten. "Norms and Nature in Translation Studies." *Incorporating Corpora: The Linguist and the Translator*, edited by Gunilla Anderman and Margaret Rogers, Multilingual Matters, 2007, pp. 49–59, <https://doi.org/10.21832/9781853599873-007>.
- . "What Happened to God and the Angels: An Exercise in Translational Stylistics." *Target. International Journal of Translation Studies*, vol. 15, no. 1, Nov. 2003, pp. 37–58, <https://doi.org/10.1075/target.15.1.03mal>.
- Mastropierro, Lucía. "Key Clusters as Indicators of Translator Style." *Target. International Journal of Translation Studies*, vol. 30, no. 2, Apr. 2018, pp. 240–59, <https://doi.org/10.1075/target.17040.mas>.
- Mathijssen, Jan Willem. *The Breach and the Observance: Theatre Retranslation as a Strategy of Artistic Differentiation, with Special Reference to Retranslations of Shakespeare's Hamlet (1777-2001)*. 2007. Utrecht University, PhD thesis.
- Mauranen, Anna, and Pekka Kujamäki, editors. "Translation Universals." *Benjamins Translation Library*, John Benjamins Publishing Company, 2004, <https://doi.org/10.1075/btl.48>.
- Meier, Frederik, et al. "Exploring the Potential of Bootstrap Consensus Networks for Large-Scale Authorship Attribution in Luxdorph's Freedom of the Press Writings." *CEUR Workshop Proceedings*, vol. 2612, 2020.
- Munday, Jeremy. "Systems in Translation: A Systemic Model for Descriptive Translation Studies." *Crosscultural Transgressions: Research Models in Translation Studies II*, Historical and Ideological Issues, 2014, pp. 76–92.
- Newmark, Peter. *A Textbook of Translation*. Prentice Hall, 1988.
- Pym, Anthony. *Method in Translation History*. St. Jerome Publishing, 1998.
- Roded, Ruth. "Gender in an Allegorical Life of Muhammad: Mahfouz's *Children of Gebelawi*." *The Muslim World*, vol. 93, no. 1, Jan. 2003, pp. 117–34, <https://doi.org/10.1111/1478-1913.00017>.
- Rodríguez Timaná, Luis Carlos, et al. "Software to Determine the Readability of Written Documents by Implementing a Variation of the Gunning Fog Index Using the Google Linguistic Corpus." *CCIS: Communications in Computer and Information Science*, vol. 1193, 2020, pp. 409–20, https://doi.org/10.1007/978-3-030-42517-3_31.
- Saldanha, Gabriela. "Translator Style: Methodological Considerations." *The Translator*, vol. 17, no. 1, Apr. 2011, pp. 25–50, <https://doi.org/10.1080/13556509.2011.10799478>.
- Shoair, M. *Awlād Hāratinā: Sīrat al-Riwāyah al-Muharramah*. Elain Publishing, 2018.
- Snell-Hornby, Mary. "Translation Studies: An Integrated Approach." *Benjamins Translation Library*, John Benjamins Publishing Company, 2006, <https://doi.org/10.1075/btl.66>.
- Solnyshkina, Marina I., et al. "Evaluating Text Complexity and Flesch-Kincaid Grade Level." *Journal of Social Studies Education Research*, vol. 8, no. 3, 2017, pp. 238–48.
- Stamatatos, Efstathios. "A Survey of Modern Authorship Attribution Methods." *Journal of the American Society for Information Science and Technology*, vol. 60, no. 3, 2009, pp. 538–56, <https://doi.org/10.1002/asi.21001>.

- Stewart, Philip. "A Tale of Two Translations." *Amazon.com*, 2000, <https://www.amazon.com/review/R2O8JG3TZ22QVX>.
- Susam-Sarajeva, Şebnem. "Multiple-Entry Visa to Travelling Theory. Retranslations of Literary and Cultural Theories." *Target*, vol. 15, no. 1, Nov. 2003, pp. 1–36. *Crossref*, <https://doi.org/10.1075/target.15.1.02sus>.
- Theroux, Peter. "Children of the Alley: A Translator's Tale." *The Massachusetts Review*, vol. 42, no. 4, 2001, pp. 666–71.
- . "Naguib Mahfouz's Children of the Alley." *The Middle East: A Guide to Politics, Economics, Society and Culture*, edited by Barry Rubin and M E Sharpe, 2012, pp. 399–401.
- Torabi, Ali G., and Seyed H. Seyedi. "A Sociological Analysis of Naguib Mahfouz's Works with a Close Look at the Novel Children of Gebelawi (Awlad Haratana)." *Rays of Criticism in Arabic and Persian*, vol. 4, no. 13, 2014, pp. 146–125.
- Van Poucke, Piet. "Ageing as a Motive for Literary Retranslation: A Survey of Case Studies on Retranslation." *Translation and Interpreting Studies*, vol. 12, no. 1, Apr. 2017, pp. 91–115, <https://doi.org/10.1075/tis.12.1.05van>.
- Venuti, Lawrence. "Retranslations: The Creation of Value." *Bucknell Review*, vol. 47, no. 1, 2004, pp. 25–38.
- . *The Translator's Invisibility: A History of Translation*. Routledge, 2004.
- Wang, Shan, et al. "Readability Is Decreasing in Language and Linguistics." *Scientometrics*, vol. 127, no. 8, July 2022, pp. 4697–729, <https://doi.org/10.1007/s11192-022-04427-1>.
- Winters, Maeve. *A Corpus-Based Study of Translator Style: Oeser's and Orth-Guttmann's German Translations of F. Scott Fitzgerald's The Beautiful and Damned*. 2005. Dublin City University, PhD thesis.
- Youdale, Robert. *Using Computers in the Translation of Literary Style: Challenges and Opportunities*. Routledge, 2019, <https://doi.org/10.4324/9780429030345>.
- Zanettin, Federico. "Corpus Methods for Descriptive Translation Studies." *Procedia - Social and Behavioral Sciences*, vol. 95, Oct. 2013, pp. 20–32, <https://doi.org/10.1016/j.sbspro.2013.10.618>.

Appendix

(a) Top 10 nouns in both translations.

Gebelawi (Stewart)		Alley (Theroux)	
man	619	man	693
alley	551	alley	668
people	497	gabal	527
gebel	494	people	432
House	440	qassem	425
qaasim	426	House	414
strongman	376	rifaa	401
rifaa	343	adham	354
adham	341	eye	351
eye	323	voice	290

(b) Top 10 adjectives in both translations.

Gebelawi (Stewart)		Alley (Theroux)	
good	289	good	282
old	143	other	137
other	119	old	131
first	102	long	109
bad	99	own	107
great	89	first	106
afraid	85	same	90
own	82	only	82
more	79	little	79
long	75	afraid	76

(c) Top 10 verbs in both translations.

Gebelawi (Stewart)		Alley (Theroux)	
be	5391	be	5378
have	1818	have	1788
say	1537	say	1430
do	1170	do	1323
go	802	go	645
come	538	come	497
see	418	see	437
look	410	know	367
know	359	look	362
ask	342	ask	356

(d) Top 10 adverbs in both translations.

Gebelawi (Stewart)		Alley (Theroux)	
not	1499	not	1418
then	493	then	411
back	309	so	311
as	275	as	285
so	256	back	267
only	203	even	245
away	176	never	213
now	146	only	203
all	146	now	195
even	143	here	174