Keywords in *Wolves of the Crescent Moon*: Thematic and Corpus Analysis

Wedad Albeyali[1] & Haneen Al-Marzouki[2], & Arwa Hasan[3]

[1] College of Language Sciences, King Saud University, Riyadh, Saudi Arabia and Technical Digital College, Technical and Vocational Training Corporation, Riyadh, Saudi Arabia

[2] Yanbu English Language and Preparatory Year Institute, Yanbu Industrial College, Yanbu, Saudi Arabia

[3] College of Language Sciences, King Saud University, Riyadh, Saudi Arabia

Correspondence: Wedad Mohammed Albeyali, Technical Digital College, Technical and Vocational Training Corporation, Riyadh, Saudi Arabia. E-mail: wedad.albeyali@hotmail.com

Abstract

Corpus methods allow for the analysis of literary texts, specifically novels, because they can objectively examine large datasets through systematic computer-based analysis. This study examined the most commonly occurring content words in the English translation of the Saudi novel *Wolves of the Crescent Moon* (WOTCM). This study implemented a mixed-methods approach utilizing the LancsBox software for a corpus-based quantitative analysis to identify the most frequently used content terms in WOTCM. The keywords within the context of the novel were further analyzed by comparing the terms to the semantic meanings that may be prompted by the translated title as well as the original title of the text through qualitative analysis of concordances. The findings revealed excessive usage of specific terms related to the titles, which may be attributable to the translator's deliberate choices or the inherent aspects of the topics that could impact the interpretation of these titles. These choices highlight the importance of paratexts, which include the titles of literary works, in determining the thematic and conceptual frameworks within the novel. The study underscores the crucial significance of corpus linguistics in comprehending translation processes and the linguistic tactics employed in literary themes and sheds light on translation choices in Saudi literary works.

Keywords: Arabic literature, content words, corpus linguistics, paratext, Saudi novel

1. Introduction

In recent decades, the integration of corpus linguistics into academia has demonstrated its relevance in language studies owing to the expansion of the digital age. The integration of stylistics with corpus linguistics has facilitated the successful merging of these two fields (Biber, 2011; Mahlberg, 2010; Wynne, 2006). The empirical characteristics of both stylistics and corpus linguistics have significantly contributed to developing their interdependent connection, leading to the formation of "corpus stylistics." Studying literary text styles can be defined as applying corpus linguistic methodologies to analyze textual meanings using empirical data (Mahlberg, 2015; Maiwald, 2011).

Corpus linguistics has proven to be highly effective in analyzing literary texts, particularly novels, attributed to its capacity to systematically examine large datasets through computer-based analysis (Mahlberg, 2015). Its popularity in the academic community is increasingly evident, notably in the analysis of English novels. One example is Wijitsopon (2013), who used corpus linguistics to investigate lexical patterns found in Jane Austen's six main novels. Similarly, Nada and Ajwan's (2023) study of *The Lord of the Flies* by William Golding employed corpus linguistics to determine the categories and semantic interpretations of modal auxiliaries by examining concordance and frequency patterns using the AntConc software tools.

Furthermore, corpus linguistics is also utilized to analyze Arabic novels translated into English. For example, Alhamshary's (2024) study investigated the construction and communication of meaning in Naguib Mahfouz's Cairo Trilogy by analyzing collocational patterns using the Sketch Engine program. In addition, Altamimi (2016) employed a corpus-driven approach to examine Humphrey Davies 'translator style by comparing his English translation of the Arabic novel *Midaq Alley* with another translator's version. These studies demonstrate the usefulness of corpus linguistics in investigating the linguistic complexities and stylistic nuances that impact the transmission of meaning within and between texts.

Yousef Al-Mohaimeed's *Wolves of the Crescent Moon* (WOTCM) has been acknowledged by many critics as a notable accomplishment in modern Saudi literature since its publication in English (Freedenberg, 2007; Lytal, 2022). Originally titled 'فخاخ الرائحة' or translated literally as Traps of Scent (TOS), Al-Mohaimeed's novel has received widespread recognition for its unique structure and bold issues, establishing it as his most prominent work so far during a time when Saudi fiction writers were reshaping current Saudi fiction standards (Al-Sharqi, 2017).

To the best of the researchers' knowledge, no study has thus far focused specifically on the linguistic analysis of WOTCM through a

corpus study. Therefore, this study aims to investigate the most frequent content words in the novel, mainly in the context of the English translation WOTCM. Further, the study aims to analyze the semantic fields of the two titles to understand their thematic and conceptual structures in-depth. Examining the most frequently used content words in the novel is crucial from a linguistic aspect. This study can reveal the translator's lexical choices, providing insights into the narrative's language texture and stylistic components. Identifying essential keywords makes it easier to recognize recurring themes, and investigating the semantic connotations of these terms helps us better comprehend the linguistic nuances of the novel. A thorough linguistic study of the novel's essential components can be obtained by exploring the semantic fields of the titles, which offers insights into the linguistically selected choices and possible symbolism incorporated within the novel's framework. Therefore, the study aims to answer the following questions:

1. What are the significantly frequent content words in the novel *Wolves of the Crescent Moon*?

2. How do the semantic fields of these two titles potentially influence the novel's interpretation?

This study tentatively hypothesizes a predominance of content words pertaining to the English title WOTCM and the Arabic TOS. It is proposed that words semantically related to the ideas "Wolves," "Religion," and "Smell" will be more prevalent in the text. This research direction could potentially shed new light on the semantic priming aspects of translated titles.

2. Literature Review

2.1 Corpora

The term "corpus" typically denotes a large compilation of authentic texts that have been preserved in a format usable by machines (McEnery & Hardie, 2012). Corpus is usually employed by linguists to refer to a compilation of authentic instances of language, ranging from a few sentences to a collection of written texts or audio recordings, which have been compiled for linguistic analysis (Hunston, 2002). Subsequently, the texts are analyzed utilizing diverse computer programs within the field of linguistics referred to as "corpus linguistics." Its application has extended to fields such as sociolinguistics (Baker, 2010), discourse analysis (Ancarno, 2020), the art of translation (Mikhailov & Cooper, 2016), literary studies (Mahlberg et al., 2019), forensic linguistics (Wright, 2021), language pedagogy (Jones & Durrant, 2010), and the analysis of social media discourse (Al-Laith et al., 2021).

2.2 Corpus Linguistics and Stylistics

Corpus linguistics and stylistics are two areas within the broader field of language studies that have shown increasing convergence, particularly with the rise of digital text analysis tools. Corpus linguistics involves the analysis of this corpus to understand patterns and structures of language use. In contrast, stylistics is the study and interpretation of texts from a linguistic perspective (Hardie & McEnery, 2010). Several publications have demonstrated the use of corpus linguistic concepts and methodologies in literature analyses. For example, Louw (1998) illustrated how a corpus can be utilized to evaluate intuitions and how symbols can be analyzed in terms of collocations. *Romeo and Juliet* has also been analyzed by examining the usage of significant vocabulary (Culpeper, 2002; Scott & Tribble, 2006). The study conducted by Adolphs and Carter (2002) explored semantic prosodies in Virginia Woolf's works. Semino and Short (2004) scrutinized the illustration of speech and thoughts. Stubbs (2005) examined several quantitative methodologies for analyzing Conrad's *Heart of Darkness*, whereas Fischer-Starcke (2006) focused on trigrams in Austen's *Persuasion*. While this list provides merely illustrative examples, it is noticeable that there are several manifestations of corpus approaches to literature. This current study, therefore, contributes to the application of corpus methods in the analysis of literary works, particularly in the exploration of its application in the genre of Saudi fiction.

2.3 Wolves of the Crescent Moon

This Saudi novel, translated into English by Anthony Calderbank, offers an overview of three significant characters who face the challenges of being marginalized and finding their true identities. The novel portrays the experiences of Turad, Tawfiq, and Nasir, who face physical disfigurement, destitution, and isolation in the demanding urban setting. The tale is structured into 16 chapters, employing diverse perspectives, temporal frameworks, and occurrences to depict the interrelated lives of the people. The narrative begins with Turad, a Bedouin who has lost his dignity, at a bus station. It also delves into Tawfiq's history as a formerly enslaved Sudanese person and Nasir's experience of being abandoned by unknown parents.

The novel has been analyzed from several angles, thoroughly examining thematic and structural aspects, evaluating cultural portrayal in translation, and investigating Saudi reactions to modernity through the novel's thematic dualities. Gohar (2015), for example, conducted a critical and comparative literary analysis of Al-Mohaimeed's novel, positioning African-American slave narrative discourses within it. He explored the relationship between slavery and liberation by delving into the novel's portrayal of the subaltern voice of the slave narrator. The study also examined the author's attempt to reconstruct thematic and aesthetic motifs hidden in African American slave narratives, including a discussion of Arabic-Islamic enslavement, a subject not well-known in contemporary Western literary canons. Gohar (2015) emphasized the connections between historical and cultural entanglements through close reading and critical analysis, shedding light on the complexity of racism and enslavement in the Arab world. The study's conclusion offered a comprehensive perspective on the work within the framework of Arabic and African-American literature, illuminating the intricate and horrifying realities of slavery and its lasting repercussions on both the individual and society.

In the realm of literary translation and cross-cultural understanding, Aldawood (2017) investigated how cultural components were portrayed in Anthony Calderbank's English translation of Al-Mohaimeed's novel through a qualitative analytical study. The study sought

to evaluate whether the translation reinforces stereotyped depictions of Arab culture and examined the influence of globalization and enhanced communication on translation tactics and cultural portrayal. The research proposed that the translator employed many tactics, including transliteration, foreignization, and domestication, to transfer cultural references and idiomatic idioms from the original Arabic text to English. The findings highlighted whether translation tactics may affect how the "Other" culture is represented, demonstrating the translator's responsibility for faithfully expressing cultural nuances and dispelling or supporting preconceptions in literary translations.

The analysis of Al-Mohaimeed's novel reveals thematic, structural, and symbolic elements, which can enhance the understanding of its literary and cultural significance as illustrated by Al-Sharqi's (2017) study. The researcher delved into how Saudis reacted to modernity by identifying and analyzing dualities in the novel, such as city versus desert, dominance versus marginalization, and nature versus culture. The study explored the use of binary opposition as a structural element in the novel to investigate the relationship between modernity and culture in Saudi Arabia, challenging previously accepted features of modern society. Al-Shargi (2017) explored characteristics, settings, and imagery to demonstrate how these binary oppositions functioned as a lens to reflect the societal impact of modernization in Saudi Arabia. The findings revealed a complicated web of dualities that bring unrelated facts into imaginative coherence, representing the alteration of traditional Saudi culture and the marginalization of specific socioeconomic groups.

2.4 Paratextual Function of the Title

Paratext refers to the supplementary components accompanying a primary text, including the title, foreword, dedication, afterword, and design elements such as cover art and typography. The variables mentioned substantially impact how readers perceive and understand a piece of work (Genette, 1997). This study aims to investigate the importance of the title within the paratextual framework, specifically its ability to direct the reader's comprehension of the text and the author's intended message. Saldanha (2014) emphasizes that various factors, including the translator's comprehension of the ST, the translation objectives, and the target audience's characteristics, influence the translation outcome. These elements are not always apparent solely through textual analysis (Saldanha, 2014).

There has been little attention given to the title of the translated WOTCM, which may have implications for the reader's interpretation of the work. The title of the text has different functions, including the appellative function of drawing readers to the work and guiding the interpretations of the novel (Nord, 2019). By changing the title from "Traps of Scent" to "Wolves of the Crescent Moon" the translated work primes the reader to certain expectations within the novel and helps provide the implication of an exotic text from a different culture. The target title of WOTCM may also focus on certain interpretative functions that are different from the source title TOS (Viezzi2011:187), which leads to highlights the themes of the desert and its wolves as well as the Islamic cultural connection through the crescent moon. A search of Goodreads reviews, which provides online written reviews of books, shows that the English title highlights the presence of wolves in the novel. For example, one reader indicated that "[. . .] the title *Wolves of the Crescent Moon* refers to both the wolves that tore Turad's ear off and Islam itself (the crescent) [sic]." Another reader indicates that the novel uses "the symbolism of the wolf" and a third indicates that the "double entendre of the title" was appreciated (*Goodreads* 2024).

These reviews indicate that the title functioned to "guide the readers' interpretation of the text" (Nord 2019: 340), and therefore the readers focused on aspects related to deserts and wolves rather than the theme of scent that may be primed by the original title TOS. To ascertain whether these themes are primed only by the title or through repetition throughout the text, a keyword analysis of content words related to the two titles may provide an objective explanation of the themes' prevalence in these readings. Although WOTM has been analyzed thematically in the studies mentioned above, this study is innovative in employing a corpus-based approach that allows for an objective investigation of these themes and whether there is an overuse of certain keywords derived from each of the two titles.

3. Methodology

This study employed a mixed-methods approach. Initially, the analysis utilized LancsBox software to conduct a corpus-based quantitative analysis to determine the most frequently used content words in the novel WOTCM**.** The keywords were analyzed qualitatively, which involved scrutinizing concordance lines to comprehend their semantic usage within the context of the novel. In addition, the frequency of these words was examined to establish their relationship with the themes of the English title WOTCM, which were classified as "Wolves," "Moon," and "Religion." As for the Arabic title TOS, the themes were categorized as "Traps" and "Scent."

To compare the most frequently used content words in the WOTCM novel with those in other Arabic novels translated by the same translator, Anthony Calderbank, a main corpus (MC) of the WOTCM novel, was compiled, comprising 31,388 tokens. Similarly, reference corpora (RC) of six different novels were created, totaling 320,750 tokens. According to Scott and Tribble (2006), keywords are significant lexical terms in a text solely due to their distinctive occurrence when compared with reference corpora of a particular appropriate type. A prevalent challenge faced by corpus compilers pertains to copyright concerns associated with the texts to be incorporated into the corpus. Accordingly, procedures similar to those used by the compiler of the King Abdulaziz City for Science and Technology (KACST) corpus were implemented. This includes restricting text distribution, download, and preview (Al-Thubaity, 2015).

After building the corpus, the most frequent content words lists were examined within the concordance lines of the novels to determine their semantic relation to the titles, a reduction procedure was then employed to exclude irrelevant words from the two translated titles, thereby enabling an in-depth analysis of frequently recurring content words related to the titles WOTCM and TOS. The removal involved eliminating figurative terms that had no direct relevance to the titles, as the aim was to focus on the literal meanings associated with the titles. In addition, items with frequencies less than two were eliminated. This thorough curation yielded a refined list of semantically relevant words.

4. Results

Theme categories were determined to compare the frequency of specific words in the MC and the RC for the titles WOTCM and TOS. The English title WOTCM was categorized into three distinct themes: "Wolves," "Moon," and "Religion." However, the Arabic title TOS was divided into two different categories: "Traps" and "Scent." Each category included all pertinent lexical variations within a particular thematic structure, considering differences associated with morphological variations. For instance, in the category labelled "Wolves," all lexicons related to wolves, such as "Wolf," "Wolves," and "Wolf's," were combined to ensure a thorough inclusion of terms that are directly relevant to the overarching theme. This approach allows for a comprehensive and all-encompassing analysis of interconnected terms within a particular thematic structure, offering a holistic view of the occurrence and utilization of words linked to each category. In addition, the lexicons in the tables below were ranked according to their log-likelihood (LL) value, which indicates whether they are overused or underused in MC compared to RC.

Table 1. The Frequency of Wolves of the Crescent Moon

Category	Lexicon	Raw Frequency (MC)	Relative Frequency (MC)	Raw Frequency (RC)	Relative Frequency (RC)	LL	%DIFF
Wolves	Wolf / Wolves / Wolf's / She-wolf / Wolfish	61	0.19	18	0.01 +	213.52	3363.06
	Sand / Sandy	48	0.15	87	0.03 +	72.61	463.80
	Wild / Wilderness	26	0.08	41	0.01 +	43.88	548.03
	Beast / Beasts / Beast's	13	0.04	8	0.00 +	36.44	1560.57
	Animal / Animals	21	0.07	33	0.01 +	35.53	550.29
	Desert / Deserts	38	0.12	112	0.03 +	34.86	246.71
	Sniff / Sniffed	6	0.02	2	0.00 +	20.39	2965.66
	Cave / Caves	7	0.02	4	0.00 +	20.17	1688.30
	Howl / Howled / Howling	12	0.04	25	0.01 +	16.06	390.51
	Blood	9	0.03	22	0.01 +	10.27	318.04
	Hyenas	2	0.01	3	0.00 +	3.50	581.26
	Kill / Kills / Killed	11	0.04	60	0.02 +	3.16	87.35
	Enemy	5	0.02	19	0.01 +	3.16	168.92
	Prey	8	0.03	39	0.01 +	3.08	109.62
	Dune	2	0.01	10	0.00 +	0.72	104.38
Moon	Moon / Moons / Moonlight / Moonlit	29	0.09	51	0.02 +	44.97	481.07
	Night / Nights	71	0.23	286	0.09 +	40.53	153.69
	Midnight	5	0.02	10	0.00 +	6.95	410.94
	Crescent	1	0.00	6	0.00 +	0.21	70.31
Religion	Hell	24	0.08	16	0.00 +	65.19	1432.83
	Hajj / Hajjis	15	0.05	5	0.00 +	50.97	2965.66
	Pilgrim / Pilgrims / Pilgrimage	10	0.03	18	0.01 +	15.21	467.72
	Heaven / Heavens	9	0.03	23	0.01 +	9.79	299.87
	God	6	0.02	136	0.04 -	4.69	-54.92
	Pray / Prayer / Praying / Prayers	6	0.02	100	0.03 -	1.57	-38.69
	Worship	2	0.01	8	0.00 +	1.16	155.47
	Allah / Allah's	9	0.03	128	0.04 -	1.01	-28.15
	Fate / Fates	7	0.02	48	0.01 +	0.88	49.03

Table 1 demonstrates the frequency of specific words within the thematic categories of "Wolves," "Moon," and "Religion," based on the English title, WOTCM, indicating significant differences in their occurrence between the MC and the RC. The LL value of words in the "Wolves" and "Moon" categories indicates an overuse in the MC compared to the RC. For instance, in the category "Wolf," the terms "Wolf/ Wolves/ Wolf's/ She-wolf/ Wolfish" occur 61 times in the MC, with a relative frequency of 0.19. Conversely, they are observed 18 times in

the RC, with a relative frequency of 0.01. After calculating the LL, it appears that these terms are significantly overused in the MC. On the other hand, based on LL, the term "Dune" is the least frequently used word, with a relative frequency of 0.01 in the MC, appearing twice, while it occurs 10 times in the RC, with a relative frequency of 0.72.

In the category of "Moon," the terms "Moon/ Moons/ Moonlight/ Moonlit" occur 29 times in the MC, with a relative frequency of 0.09, whereas they appear 51 times in the RC, with a relative frequency of 0.02, in which the LL value demonstrates their overuse in the MC. In contrast, according to the LL value, the least frequently used term is "Crescent," which appears only once in the MC with a relative frequency of 0.00, and six times in the RC with a relative frequency of 0.21.

Regarding the "Religion" category, the LL value presents extensive use of most words in the MC relative to the RC. Among these terms, "Hell" illustrates the highest frequency, appearing 24 times with a relative frequency of 0.08, whereas the terms "Fate/ Fates" demonstrate the lowest frequency, occurring 7 times in the MC, with a relative frequency of 0.02. Conversely, based on the LL value, the terms "Allah/Allah's," "Pray/ Prayer/ Praying/ Prayers," and "God" are underused in the MC compared to the RC. As an illustration, the terms "Allah/Allah's" appear 9 times in the MC, with a relative frequency of 0.03, yet 128 times in the RC, with a relative frequency of 0.04.

Table 2. The Frequency of Traps of Scent

Category	Lexicon	Raw Frequency (MC)	Relative Frequency (MC)	Raw Frequency (RC)	Relative Frequency (RC)	LL	%DIFF
Traps	Tie / Tied	19	0.06	63	0.02 +	14.85	208.19
	Slave / Slaves / Slave's	32	0.10	155	0.05 +	12.50	110.97
	Wrapped	13	0.04	44	0.01 +	9.86	201.92
	Catch / Caught	11	0.04	42	0.01 +	6.90	167.64
	Snatched	6	0.02	15	0.00 +	6.68	308.75
	Servant / Servants	12	0.04	59	0.02 +	4.53	107.84
	Captured	2	0.01	3	0.00 +	3.50	581.26
	Prey	8	0.03	39	0.01 +	3.08	109.62
	Secret / Secrets / Secretly	15	0.05	94	0.03 +	2.75	63.07
	Trap / Trapped	6	0.02	29	0.01 +	2.36	111.42
	Death / Die / Died	14	0.04	165	0.05 -	0.27	-13.29
Scent	Smell / Smelled / Smells	40	0.13	83	0.03 +	53.75	392.48
	Sniff / Sniffed	6	0.02	2	0.00 +	20.39	2965.66
	Scent	9	0.03	17	0.01 +	13.15	441.00
	Flowers	11	0.04	30	0.01 +	11.10	274.69
	Nose	11	0.04	47	0.01 +	5.62	139.17
	Aroma	2	0.01	4	0.00 +	2.78	410.94
	Stench	1	0.00	2	0.00 +	1.39	410.94
	Odor	1	0.00	2	0.00 +	1.39	410.94
	Perfume	7	0.02	52	0.02 +	0.58	37.56

According to the Arabic title, TOS, Table 2 displays the frequency of particular words within the thematic categories of "Traps" and "Scent." By calculating the LL value, the results highlight significant variations in the occurrence of these terms between the MC and the RC. In the "Traps" category, the terms "Tie/ Tied" are found 19 times in the MC, presenting a relative frequency of 0.06, while, in the RC, these terms are used 63 times, with a relative frequency of 0.02, in which the LL value suggests that they are the most often utilized terms in the MC. However, the LL value reveals that the terms "Trap/ Trapped" are the least commonly utilized, appearing only 6 times, with a relative frequency of 0.02 in the MC. Considering the underused terms in the MC, the term "Death" occurs 14 times, with a relative frequency of 0.04, yet in the RC, it appears 165 times, with a relative frequency of 0.05.

In the "Traps" category, the LL value indicates that all the terms demonstrate a tendency towards overuse. Notably, the terms "Smell/ Smelled/ Smells" demonstrate the highest frequency in the MC compared to the RC. Specifically, they appear 40 times in the MC, showing a relative frequency of 0.13, while, in the RC, they occur 83 times, displaying a relative frequency of 0.03. In contrast, the term "Perfume" is the least frequently used in which it occurs 7 times in the MC, with a relative frequency of 0.02, whereas it is seen 52 times in the RC, with a relative frequency of 0.02.

5. Discussion and Conclusion

The above findings shed light on two key aspects related to the research questions tackled in this study: (1) the significantly frequent content words related to the two titles (2) and how the semantic fields of these two titles potentially influence the novel's interpretation. This section will address each of these points simultaneously. As previously mentioned, this study hypothesized an overuse of content words related to each translated title. Therefore, given that the titles are WOTCM and TOS, the hypothesis posited that the semantic meanings were related to the words "Wolves," "Religion," and "Smell."

The analysis of keywords utilized LL as a method for determining significance of frequency, where frequency is considered significant when the log-likelihood calculation of word occurrence across the two corpora is greater than 6.63, which indicates that the probability of the difference occurring by chance is p<0.01. The results section demonstrated that, within the English title, WOTCM, the term "Wolves" appeared as the most frequent word in the corresponding semantic category. As shown in Table 1, this overuse is indicated by the LL value 213.52. Within the categories "Moon" and "Religion," the words "Moon" and "Hell" respectively manifested as keywords with the respective values 44.97 and 69.15. Notably, the religious lexicon, including "Allah" and "God," was utilized less frequently than anticipated, with the respective values 1.1 and 4.69, which shows an underuse in comparison to the reference corpora. This underuse could be considered in line with the themes of the novel that focused on the clash between modernity and tradition, highlighting an emphasis on the thematic significance of the lone wolf in a society moving away from the solace of traditional religious connection with God towards a more secular modernity.

In the analysis of the Arabic title TOS, which contains the semantic categories of "Traps" and "Scent," the word "Tie" emerged as the most frequently occurring term in the "Traps" category, meanwhile "Smell" was identified as the predominant term in the "Scent" category.

The frequent recurrence of specific terms in a translated work might be correlated with various factors. This could represent the translator's strategic choice to achieve literary and stylistic goals, such as theme reinforcement. By employing this technique of repetition, the translator adeptly guides the reader's focus towards the central themes and motifs, thereby fostering a stronger connection with the core of the novel. Additionally, in the Arabic text of TOS, the overused term "Smell" occurs 10 times within one passage in the novel, in which the character Turad reflects on his trapped situation (Almohaimeed, 2003: 85-86). This overuse in a particular passage illustrates the internal foregrounding technique where there is what is referred to as "internal deviation" with the high recurrence of the word in contrast to the norm established by the rest of the text (e.g., Halliday, 2020; Levin, 1965). This may not only distinguish this passage from the remainder of the text, but also signify a stylistic decision to draw attention to its thematic significance in the novel. This passage can be considered a critical point in the novel, where all three protagonists, Turad, Nahar, and Amm Tawfiq, converge, offering a distinct portrayal of their interconnection as expressly elucidated by the author. The commonality throughout these stories is not the wolf or religion, but rather the notion of being trapped by scent, highlighting that this passage constitutes the climax of the narrative, reinforcing the central motif highlighted in the original title TOS, but may not be as salient in readings under the title of WOTCM.

A further aspect of this recurring pattern related to both titles could be interpreted as stemming from the inherent nature of the content. A review of Goodreads reviews has revealed that translation discrepancies have shifted the focus of readers of the English title toward the term "Wolf," which is influenced not only by the title but also by the frequent use of the word "Wolf" in the text (Goodreads, 2024). For instance, concerning the title WOTCM, the frequent use of "Wolves" corresponded to the expectations the title sets. Concurrently, with TOS, the anticipation of encountering terms semantically related to "Scent" and "Smell" is logically met only in the climax of the novel.

The frequent use of specific terms throughout the novel may be related to the interpretation of the two titles. In any literary work, the title is the first thing the reader encounters, and thus may influence the reader's response to a text. Lodge supports this notion, stating that "[t]he title of a novel is part of the text - the first part of it, in fact, that we encounter - and therefore has considerable power to attract and condition the reader's attention" (1992, p.193). Therefore, the paratext of the title has a function in encouraging certain responses to the text, which is also shown in the current study's results with the correlation of the keywords to each title. The translator seems to have modified the title from TOS to WOTCM to establish a stronger connection between the content and Saudi culture, simultaneously catering to the Western readers' preferences. Wolves are typically attributed to Arabian deserts, whereas the crescent moon is frequently connected to Islam. However, Western readers may not find the title TOS to be relevant to the narrative because they may lack knowledge about the cultural references in the Arabic setting. Accordingly, the research findings indicate that the terms associated with both titles can be considered as keywords in the corpora. This may indicate a purposeful choice to draw attention to wolves and religious motifs, which are regarded as significant in the themes of the narrative.

Within the realm of literary analysis, the studies by Al-Sharqi (2017) and Gohar (2015) focused on thematic elements, exploring how the content of novels reflects societal changes and underlying conflicts. Al-Sharqi (2017) specifically examined societal responses to modernity, while Gohar (2015) drew parallels with African-American slave narratives. In contrast, this study adopted a linguistic approach, focusing on how choices in translation, particularly of titles, impact thematic interpretation. The underuse of the aforementioned religious lexicon reinforces the discussion of how modern society casts the three main characters into isolation, and away from religious solace. Whereas Al-Sharqi's (2017) analysis shows the binary oppositions of tradition versus modernity, the corpus analysis demonstrated more overuse with words like "Hell" indicating the departure from hope for the three protagonists, indicating that this clash has led to a more depressing outcome and highlights the tragic effect of modernity on a traditional landscape. WOTCM therefore presents a harsh modernity that enslaves its characters, and this enslavement is also shown in the overuse of keywords such as "Slave" that occurred with a log likelihood of 12.5, which also shed light on the underlying slave narrative that was analyzed by Gohar (2015). Additionally, this study aligned with Aldawood's (2017) research in examining translation aspects, although with different emphases. Aldawood (2017) analyzed broader translation processes and their influence on cultural representation, including how they can perpetuate stereotypes. This study, however, delved into how specific lexical choices in translated titles influence thematic interpretation. The findings show that although the title choice of WOTCM seemed to have the appellative function of drawing readers to cultural elements in the novel, they are in line with the keyword analysis that highlights the presence of the words related to "Wolf." Therefore, the

introduction of the wolf in the title of the novel can highlight and reinforce the theme of loneliness in the modern setting as demonstrated in the literature.

The corpus analysis in this research emphasizes the impact of linguistic choices in translated titles and their role in shaping thematic interpretations. This study demonstrates that specific words in translations can significantly alter readers' perspectives, providing a sophisticated understanding of the nuanced role of language in shaping a text's thematic reception. It focuses on the effect of two translated titles, setting it apart from other research that emphasizes broader thematic or cultural analysis. Therefore, while all the related studies mentioned in this study aimed to enhance our understanding of the reception of the narrative across various contexts, this research uniquely underscores the critical role of language in translation, particularly in how it influences thematic interpretation through title selection.

The overarching goal of this study was to determine the potential impact of frequently used content words on readers' interpretations of the two titles. The results revealed the overuse of particular terms pertaining to the titles, which could be attributed to the translator's strategic choices or intrinsic facets of the themes that might influence the interpretation of these titles.

Although this study has made significant contributions to the field of corpus linguistics, it is crucial to acknowledge its limitations, which can serve as a guide for future research. This study concentrated exclusively on translated versions of novels, analyzing solely the English corpus. To conduct a more thorough linguistic analysis, future studies could broaden their scope by creating and comparing corpora in both the original Arabic text and its translated English version. In addition, the present study was restricted to analyzing translations conducted by Anthony Calderbank. Future research could expand the investigation to encompass comparisons between several novels by Yousef Al-Mohaimeed or between his works and other Saudi literary texts. Furthermore, although this analysis primarily examined the explicit meanings of frequently used words in the novels, future research could investigate the figurative language usage. This would enhance the comprehension of linguistic mechanisms that determine the stylistic and thematic terms in literary texts.

The results of this study have significant implications for the way translations are conducted, the advancement of tools in computational linguistics, and cross-linguistic comparisons. Specifically, by revealing the substantial influence of word selection on the interpretation of themes, this research could potentially affect how translators and marketers approach the translation of literary titles. Furthermore, this research highlights the collaborative nature of corpus studies, particularly the importance of cross-language comparisons. These comparisons can provide invaluable insights into how diverse linguistic communities perceive translated literature. This inclusive approach has the potential to foster the creation of bilingual or multilingual corpora, thereby enhancing the capacity for comparative analyses. By addressing these implications, the study emphasizes the crucial significance of corpus linguistics in enhancing our comprehension of translation processes and the diverse linguistic tactics utilized in literature. Corpus linguistics, which concentrates on extensive language data and statistical analysis, provides a distinctive perspective on translation, allowing us to study and measure patterns of language and thematic interpretation.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).

Copyrights

References

Adolphs, S., & Carter, R. (2002). Points of view and semantic prosodies in Virginia Woolf's *To the Lighthouse*. *Poetica*, *58,* 7-20.

Aldawood, S. (2017). Representing the other in the translation of culture-specific elements in literary texts: Wolves of the Crescent Moon as a case study. *International Journal of Language and Linguistics, 4*(2), 29-40. Retrieved from https://bit.ly/4dluDbR

Alhamshary, Y. (2024). Digital readings on collocational networks in Mahfouz's translated "trilogy" from frequencies to creating contextual cultural meaning. *Egyptian Journal of Linguistics and Translation*, *12*(1), 80-153. https://doi.org/10.21608/ejlt.2023.241124.1043

Al-Laith, A., Shahbaz, M., Alaskar, H. F., & Rehmat, A. (2021). Arasencorpus: A semi-supervised approach for sentiment annotation of a large Arabic text corpus. *Applied Sciences*, *11*(5), 2434. https://doi.org/10.3390/app11052434

Al-Mohaimeed, Y. (2003). *Fikhakh Al-Raeha* [Traps of Scent]. Riyadh AlRayes Books and Publishing.

Al-Mohaimeed, Y. (2010). *Wolves of the crescent moon* (A. Calderbank, Trans.). The American University in Cairo Press.

Al-Sharqi, L. (2017). Patterns of binary oppositions in Yousef Al-Mohaimeed's Wolves of the Crescent Moon. *Advances in Language and Literary Studies*, *8*(5), 1-8. https://doi.org/10.7575/aiac.alls.v.8n.5p.1

Altamimi, B. A. (2016). *A corpus-driven investigation of translator style: A study of Humphrey Davies' Arabic-English translations of Midaq Alley and the Yacoubian Building* [Doctoral dissertation, University of Leeds].

Al-Thubaity, A. O. (2015). A 700m+ Arabic corpus: KACST Arabic Corpus Design and construction. *Language Resources and Evaluation*, *49*(3), 721-751. https://doi.org/10.1007/s10579-014-9284-1

Ancarno, C. (2020). Corpus-assisted discourse studies. In A. de Fina & A. Georgakopoulou (Eds.), *The Cambridge handbook of discourse studies* (pp. 165-185). Cambridge University Press.

Baker, P. (2010). *Sociolinguistics and Corpus Linguistics.* Edinburgh University Press.

Biber, D. (2011). Corpus linguistics and the study of literature. *The Future of Scientific Studies in Literature*, *1*(1), 15-23. https://doi.org/10.1075/ssol.1.1.02bib

Culpeper, J. (2002). Computers, language and characterisation: an analysis of six characters in Romeo and Juliet. In U. Melander-Marttala, C. Ostman, & M. Kyto (Eds.), *Conversation in Life and in Literature: Papers from the ASLA Symposium* (pp. 11-30). Association Suedoise de Linguistique Appliquee.

Fischer-Starcke, B. (2006). The phraseology of Jane Austen's Persuasion: Phraseological units as carriers of meaning. *ICAME Journal, 30.* 87-104. Retrieved from https://bit.ly/4diCozr

Freedenberg, H. (2007, December). A portrait of real life in Saudi Arabia. *BookPage*. Retrieved from https://www.bookpage.com/reviews/5561-yousef-al-mohaimeed-portrait-real-life-saudi-arabia-fiction/

Genette, G. (1997). *Paratexts: Thresholds of interpretation*. Cambridge University Press.

Gohar, S. (2015). Positioning African American slave narrative discourses in Wolves of the Crescent Moon. *Primerjalna Knjizevnost*, *38*(3), 59-81. Retrieved from https://bit.ly/3XZ35ED

Goodreads. (2024).Wolves of the Crescent Moon Reviews. *Goodreads*. Retrieved from https://www.goodreads.com/book/show/2052875.Wolves_of_the_Crescent_Moon

Halliday, M. A. K. (2020). Linguistic function and literary style: An inquiry into the language of William Golding's the inheritors. *The Language and Literature Reader,* 19-28. https://doi.org/10.4324/9781003060789-4

Hardie, A., & McEnery, T. (2010). On two traditions in corpus linguistics, and what they have in common. *International Journal of Corpus Linguistics*, *15*(3), 384-394. https://doi.org/10.1075/ijcl.15.3.09har

Hunston, S. (2002). *Corpora in Applied Linguistics.* Cambridge University Press.

Jones, M., & Durrant, P. (2010). What can a corpus tell us about vocabulary teaching materials. In A. O'Keeffe & M. McCarthy (Eds*.), The Routledge Handbook of Corpus Linguistics* (pp. 387-400). Routledge.

Levin, S. R. (1965). Internal and external deviation in poetry. *WORD*, *21*(2), 225-237. https://doi.org/10.1080/00437956.1965.11435425

Lodge, D. (1992). *The art of fiction*. Penguin Books.

Louw, B. (1998). The Role of Corpora in Critical Literary Appreciation. In A. Wichmann & S. Fligelstone (Eds.), *Teaching and Language Corpora* (pp. 240-251). Routledge.

Lytal, B. (2022, December). The First Great Saudi Novel. The New York Sun. Retrieved from https://www.nysun.com/article/arts-first-great-saudi-novel

Mahlberg, M. (2010). Corpus linguistics and the study of nineteenth-century fiction. *Journal of Victorian Culture*, *15*(2), 292-298. https://doi.org/10.1080/13555502.2010.491667

Mahlberg, M. (2015). Literary style and literary texts. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 346-361). Cambridge University Press.

Mahlberg, M., Wiegand, V., Stockwell, P., & Hennessey, A. (2019). Speech-bundles in the 19th-century English novel. *Language and Literature: International Journal of Stylistics*, *28*(4), 326-353. https://doi.org/10.1177/0963947019886754

Maiwald, P. (2011). Exploring a Corpus of George MacDonald's Fiction. *North Wind: A Journal of George MacDonald Studies, 30*(5), 50-84. https://bit.ly/3zAhopS

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Mikhailov, M., & Cooper, R. (2016). *Corpus linguistics for translation and Contrastive Studies: A guide for research*. Routledge.

Nada, A., & Ajwan, A. (2023). A corpus-based analysis of Modal Auxiliaries of William Golding's novel "The lord of flies." *Journal International of Lingua and Technology, 2(*2), 90-101. https://doi.org/10.55849/jiltech.v2i2.461

Nord, C. (2019). Paving the way to the text: Forms and Functions of Book Titles in Translation. *Russian Journal of Linguistics, 23*(2), 328-343. https://doi.org/10.22363/2312-9182-2019-23-2-328-343

Saldanha, G. (2014). Style in, and of, translation. In S. Bermann, & C. Porter (Eds.), *A Companion to Translation Studies* (pp. 95-106). Wiley Blackwell.

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. John Benjamins Publishing.

Semino, E., & Short, M. (2004). *Corpus stylistics: Speech, writing and thought presentation in a corpus of English writing*. Routledge.

Stubbs, M. (2005). Conrad in the computer: Examples of quantitative stylistic methods. *Language and Literature: International Journal of Stylistics*, *14*(1), 5-24. https://doi.org/10.1177/0963947005048873

Viezzi, M. (2011). The translation of book titles: theoretical and practical aspects. *Beyond Borders: Translation Moving Languages, Literatures and Cultures*. Berlin: Frank & Timme, pp.183-195.

Wijitsopon, R. (2013). A corpus-based study of the style in Jane Austen's novels. *MANUSYA*, *16*(1), 41-64. https://doi.org/10.1163/26659077-01601003

Wright, D. (2021). Corpus approaches to forensic linguistics. In M. Coulthard, A. May & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 611-627). Routledge.

Wynne, M. (2006). Stylistics: Corpus approaches. *Encyclopedia of Language & Linguistics*, 223-226. https://doi.org/10.1016/b0-08-044854-2/00553-8