A Computational Approach to Decode the Pragma-Stylistic Meanings in Narrative Discourse

Ayman Farid Khafaga[1], Iman El-Nabawi Abdel Wahed Shaalan[2]

College of Science & Humanities, Prince Sattam bin Abdulaziz University, Saudi Arabia[1, 2]
Faculty of Arts &Humanities, Suez Canal University, Egypt[1]
College of Education, Al-Azhar University, Cairo, Egypt[2]

Abstract—This paper presents a computer-based frequency distribution analysis to decode the pragma-stylistic meanings in one of the narrative discourse represented by Orwell's dystopian novel Animal Farm. The main objective of the paper is to explore the extent to which computer software contribute to the linguistic analysis of texts. The paper uses the variable of frequency distribution analysis (FDA) generated by concordance software to decode the pragmatic and stylistic significance beyond the mere linguistic expressions employed by the writer in the selected data. Some words were selected to undergo a frequency distribution analysis so as to highlight their pragmatic and linguistic weight which, in turn, helps arrive at a comprehensive understanding of the thematic message intended by the writer. The paper is grounded on one analytical strand: Frequency distribution analysis conducted by concordance. Results reveal that applying a frequency distribution analysis to the linguistic analysis of large data fictional texts serves to (i) identify the various types of discourse in these texts; (ii) create a thematic categorization that is based on the frequency distribution analysis of specific words in texts; and (iii) indicate that not only high frequency words are indicative in the production of particular pragmatic and stylistic meanings in discourse, but also low frequency words are highly indicative in this regard. These results accentuate a further general finding that computer software contribute significantly to the linguistic analysis of texts, particularly those pertaining to literature. The paper recommends further and intensive incorporation of computer and CALL (computer-assisted language learning) software in teaching and learning literary texts in EFL (English as a foreign language) settings.

Keywords—Frequency distribution analysis; narrative discourse; pragma-stylistic meanings; thematic categorization

I. Introduction

For a long time, the use of computer-aided text analysis software has proven useful and contributive to the linguistic investigation of texts, particularly literary texts [1]. The reason why computer-aided text analysis (henceforth, CATA) tools are positively constructive in the analysis of literary genres lies in the fact that this type of texts always abound in large amount of data, i.e. a huge and vast number of lexis that it would be difficult to be analyzed without the help of computer [2], [3]. This paper, therefore, attempts to explore the pragma-stylistic meanings pertaining to George Orwell's Animal Farm, by using a frequency distribution analysis (FDA), which is considered one of the variables of CATA, to generate the occurrences of particular selected words, which, in turn, helps decode the various pragmatic and stylistic meanings pertaining

to the text under investigation. The employment of CATA facilitates the whole process of readership and text reception on the part of readers, on the one hand, and helps clarify the different pragmatic and stylistic meanings encoded in the specific usage of each linguistic expression, on the other [4]. Linguistic expressions, in the context of this paper, include the different linguistic units, such as the word, the phrase, and the sentence. However, the analytical focus here will be on the level of the word. That is, some particular words are selected from the novel under investigation to undergo a frequency distribution analysis, by generating the total occurrences they have in the text, and the extent to which their frequency influences the general message of the text, the pragmatic meanings beyond the semantic proposition of the linguistic expressions, and the stylistic features adopted by the writer and textually reflected in the selected text.

A. Research Significance

The significance of this paper lies in its attempt to highlight the effective work of computer software in the linguistic analysis of large data texts, particularly fictional texts, by demonstrating the extent to which the pragmatic and stylistic meanings of some words can be revealed by the application of these computer software, such as the frequency distribution analysis adopted in this paper. This serves to emphasize the importance of applying and using the different computer software in the linguistic and textual analysis of texts. It is also anticipated to contribute to EFL (English as a foreign language) settings, computational linguistics, and corpus linguistics studies.

B. Research Questions

The current study attempts to answer the following research questions:

1) To what extent does a computer-based frequency analysis contribute to clarifying the pragma-stylistic meanings pertaining to particular words in the selected novel?

2) How does a frequency distribution analysis communicate specific pragmatic and stylistic meanings in the novel under investigation?

3) How do high frequency words mirror the thematic and discursive nature of the analyzed text?

4) Does the use and application of a frequency distribution analysis facilitate an intelligible perception and understanding of various types of texts?

C. Research Objectives

By applying a frequency distribution analysis to a number of selected words from the novel at hand, the study tries to achieve the following research objectives:

1) To explore the extent to which the use and application of computer software contribute significantly to the linguistic analysis of large data texts, particularly the fictional ones, as is the case with the novel under investigation.

2) To highlight the importance of incorporating the various computer software into the readership process of fictional texts.

3) To demonstrate the significance of applying computer software to the general understanding of the pragmatic, stylistic and thematic messages of literary texts.

The remainder of this article is structured as follows. Section II presents the literature review pertinent to the current study, as well as the related literature and the previous studies that addressed the same topic. Section III offers the methodology of the study, in which the selected data is described and the adopted analytical procedures are provided. Section IV demonstrates the analysis and results of the paper. Section V is dedicated to the discussion of the obtained results. Section VI is the conclusion, wherein some recommendations for future research are offered.

II. LITERATURE REVIEW

This section presents the review of literature as well as some previous studies that are relevant to the current study.

A. Computer-Aided Text Analysis (CATA)

The computer-aided text analysis offers a variety of analytical variables that can be employed in the analysis of the different types of texts [5], [6]. Various types of computer software are used in the linguistic analysis of texts, as they provide the analysts with different analytical options that help in deciphering the meanings beyond the linguistic expressions [7], [8]. According to [9], before the emergence and applications of modern technologies in teaching and learning, particularly in EFL settings, the analysis of large data texts such as the literary ones as well as the process of teaching them was very difficult. This is because the teacher and the student alike had to read the whole text in order to arrive at the meaning s/he targets, or what is called the thematic message of the text. The traditional way of reading literary texts, their reception on the part of students and their presentation on the part of teachers constituted an academic burden on both parties [10]. Furthermore, [11] emphasizes that the majority of results pertaining to the traditional analysis of literary texts are inaccurate. There is always some sort of mistake, specifically in terms of approaching the number of occurrences a given lexical item or a linguistic expression has in text. Traditional way of analyzing literary texts was a big problem, as it required much more effort and time than is the case if such analyses were conducted by means of CATA, with its different analytical options and software [12].

According to [13], the use of the computer and computational linguistics work makes it possible to process, access and examine large data for a diversity of purposes, and to investigate questions which could not plausibly be answered if the analysis was carried manually. He maintains that computer software provide various applications in automated indexing, classification, concordance, content analysis, thematic categorization and syntactic analysis. Furthermore, these computational software provide information on the company words keep in a corpus, and can also display a variety of senses of a lexical item type.

It is worth mentioning that CATA offers a number of analytical variables that contribute effectively to the process of analyzing large data texts in general, and literary genres in particular [14]. One of these variables is the frequency distribution analysis (FDA), which will analytically be applied in the current paper. An FDA serves to show the different frequencies and occurrences a given searched item has in a text [15]. One advantage of this variable is its ability to offer accurate, credible, authentic and concise results that rest beyond any supposed analytical vulnerability. According to [15], this high level of verification, credibility and authenticity attained by the application of FDA reflects the relevance of applying computer software to the analysis of the different types of texts.

Another analytical variable provided by CATA is what is called Key Word in Context (KWIC) [16]. This variable functions to offer the contextual environment in which the searched word occurs, which, in turn, serves to extend the pragmatic and stylistic purposes beyond the surface usage of this word. For [17], through KWIC, one can have a clear and credible picture of the preceding and subsequent words in company of the searched item. These preceding and subsequent contexts add more clarification to the pragmatic power of the linguistic expression, as well as to the way this single linguistic item contributes to the interpretation of the whole text.

According to [18], Content Analysis (CA) is another analytical variable that can be generated by CATA. This variable constitutes the process of analyzing the content produced, reproduced and maintained by any given lexical item. Significantly, by means of content analysis, one can delve into both the semantic compatibility of words; that is, the different semantic propositions these words communicate in the analyzed text and the pragmatic interpretation attributed to these words with their specific usages in texts.

A further variable offered by CATA in the context of this paper is related to the Thematic Categorization of discourse types [19]. This analytical option allows the analyst to categorize the various words with their frequencies with the themes they address in texts. Thematic distribution, therefore, is closely related to text clustering, as it identifies the different themes according to the number of frequencies words have. Crucially, this analytical variable operates in combination with KWIC; that is, it targets the thematic distribution of any searched word according to its contextual environment.

The last processing variable generated by CATA in this study is the variable of Lemmatization [20], a processing option which provides counts of lemmas (sets of grammatical words having the same stem and / or meaning and belonging to

the same major word classes, differing only in inflection and/or spelling. For [21], lemmatization serves to offer a classification of all the identical or related forms of a word under a common headword, which further functions to give a clear picture to the interpretative atmosphere of the analyzed text.

Significantly, all the aforementioned processing options can be generated by concordance, a computer software that allows researchers to access and process large data texts to produce the occurrences of particular tokens. According to Sinclair [22], the process of producing and accessing word indexes and concordances is the most obvious, conceivable and plausible application of the computer software in literary research. He emphasizes that such an automated approach to text analysis sets the analytical foundation for theoretical, empirical and analytical decisions on the various linguistic aspects of vocabulary expression, morphological, syntactic, and semantic dimensions of contained data, and the presentation of lexical and syntactic items and collocations of both high and low frequency tokens [23].

B. Previous Studies

Much previous research has been conducted on the application of CATA software into the linguistic investigation of texts. One study was presented by [24], who employed a computational approach that is based on a frequency distribution analysis to decode the extremist ideologies in the discourse of ISIS (Islamic State in Iraq and Syria). The study utilized the program of concordance to arrive at the total frequencies of specific words and collocations that help in the ultimate analysis of the discourse of ISIS. This study concluded that computer software are very effective in the analysis of all types of texts, as they offer accurate and credible results upon which discourse analysts can predicate their thematic and ideological investigation.

Another study conducted by [25] provided a computer-aided text analysis within courtroom discourse. This study employed CATA to explore the persuasive strategies used in the legal discourse of attorneys in the trial of Moussaoui that occupied the world and public opinion during the 1990s. The study clarified that CATA software are effective in deciphering the various persuasive tools employed by the two attorneys of the trial. The study further demonstrated the extent to which the number of occurrences of specific expressions within particular contexts in courtroom facilitates the process of persuasion on the part of the attorneys, either during their conversational turns with the allocated judge of the court or with testimonies of the witnesses. The study recommends the application of CATA software to the analysis of the different legal texts.

A third study by [26] investigated the ideological agency exercised by speakers over their recipients. This study uses concordance software in the analysis of the data to show the frequency distribution of the function words that indicate agency. The study also clarified that concordance proves useful in demonstrating the indicative occurrences out of the total frequency of each lexical token. This study concluded that the concept of ideological agency is better revealed via the application of computer software manifested in concordance.

A further relevant study was the one conducted by [27], in which they explored the extent to which CALL (computer-assisted language learning) software is effective in the EFL contexts. This study is entirely based on testing the effectiveness of using the two computer programs of Snagit$_{TM}$ and Screencast on acquiring the skill of reading. The study revealed that the application of the two computer software serves to improve the academic level of students, by fostering the linguistic skills pertinent to the acquisition of the skill of reading. The study also reported that such technological incorporation into EFL course functions to develop not only the linguistic competence of EFL students, but also their communicative skills. This study recommended the application of the different CALL software to the different EFL courses, as they facilitate the process of teaching and learning on the part of both teachers and students.

The previous studies so far have employed CATA software into the linguistic analysis of texts. Some of these studies focused on fictional texts, whereas other studies have presented discussions on legal texts and EFL settings. One observation concerning related literature is that it did not use CATA software within the scope of pragmatics and stylistics; that is, none of the previous studies has employed CATA software to explore the different pragmatic or stylistic purposes in discourse. This last point is the core concern of the current study, which constitutes the research gap attempted to be addressed in this article. The current study, therefore, tries to fulfill in part this research gap, by showing the effectiveness of using a frequency distribution analysis (FDA) as an indicator of discourse type and thematic categorization, as well as an analytical identifier of both the indicative and/or non-indicative occurrences in a corpus.

III. METHODOLOGY

This section presents the methodology of the study which constitutes data collection and description, the procedures adopted in the analysis of the selected data, and the rationale beyond the study.

A. Data: Collection and Description

The data in this paper comprises one literary text written by George Orwell: Animal Farm. A number of words from the novel were selected to undergo a frequency distribution analysis, which in turn clarified the high frequency words and the low frequency words and the extent to which both groups of words are indicative in reflecting the pragma-stylistic meanings pertaining to the selected novel. The selected words revolve around the discourse types of equality and inequality; the themes of oppression, rebellion and violence; and the point of view of the writer. Clarifying the way these concepts were perceived by means of the application of an FDA serves to mirror both the pragmatic and stylistic purposes targeted beyond their usage in the novel.

B. Research Procedures

Three procedural stages were adopted in this research. First, the use of the computer software adopted here has involved the preparation of the selected work by scanning and storing it electronically in order to be ready for computational analysis.

Second, the selected words were highlighted to undergo a frequency distribution analysis to generate their frequency of occurrences in the text under investigation. This stage was followed by content and thematic categorization analysis, wherein a connection has been made between the occurrences of each selected word and its significance to the pragmatics and stylistics of the novel as a whole. The third stage constitutes the interpretation and explanation of the results, as well as to relate the obtained results with the pragma-stylistic meanings communicated in the novel at hand.

C. Rationale of the Study

There are three reasons that constitute the rationale beyond the selection of Animal Farm: first, the novel has two different types of discourse: the discourse of equality and the discourse of inequality. Thus the selection of some words to undergo a frequency distribution analysis is relevant to identify the type of discourse. Second, the novel also abounds in themes that can further be decoded and categorized by the frequency analysis of a group of selected words. Third, the novel witnesses a number of words that are highly indicative in the production of the total interpretation of its incidents despite the fact that some of these words are very low in frequency.

IV. ANALYSIS AND RESULTS

A. FDA as Indicator of Discourse Type

The application of an FDA serves to demonstrate the type of discourse stylistically communicated by the writer of the novel. The number of occurrences of some particular words mirrors the type of discourse in which these words occur and address. In Animal Farm, there are two types of discourse: the discourse of equality and the discourse of inequality. Consider the following table.

TABLE I.   FREQUENCY DISTRIBUTION ANALYSIS TO SHOW TYPE OF DISCOURSE

Lexical item	Total Frequency	Indicative Occurrences
comrades	55	23
rebellion	29	2
wisdom	2	2
man	21	11
remains	1	1
miserable	5	3
slavery	3	1
free	7	1
leader	8	5
sacrifice	4	2
cruelty	2	1
criminal	1	1
traitor	3	2
laborious	3	1
tactics	2	2
comrade	41	13
equal	8	2
rich	3	1
brothers	1	1
percent	3	3
welfare	2	1
necessary	13	1
nonsense	1	1
nothing	24	1
agent	3	3
maneuver	2	1
get rid of	2	1
let us	12	2
true	7	1
enemy	12	12
friendship	1	1
friend	4	2
equality	2	2

Table I shows a number of words with the occurrences they have in the novel. The table demonstrates that there are some words pertinent to the discourse of equality, including equality, friendship, comrades, rebellion, wisdom, brother, and free. The associative meaning of closeness, brotherhood, cooperation and solidarity these words carry is an indication that they are related to the discourse of equality. It is also obvious from the table that there are some words among this group that have high frequency (e.g., rebellion, comrades), and other words that have low frequency (e.g., equality); however, the indication is that the word, regardless of its number of occurrences, may be very indicative in carrying the meaning of a specific type of discourse. In the same vein, Table I displays some words that can be perceived as indicators of an inequality discourse. Words such as man, miserable, cruelty, traitor, criminal, enemy, and remains also communicate the connotative meanings of oppression, domination and inequality. Again, the low frequency words of this group accentuate the fact that the words may be very low in frequency (e.g., slavery, laborious) but highly indicative in carrying the meaning of the discourse targeted by the writer.

B. FDA as Indicator of Thematic Categorization

The FDA can also be an indicator that clarifies the categorization of themes in texts. In the novel under investigation, one can identify a number of themes, such as violence, rebellion and oppression. These themes can be decoded by means of a frequency analysis that classifies the different words addressing a particular theme. Consider the following table.

TABLE II.   FREQUENCY DISTRIBUTION ANALYSIS OF THE WORDS CONVEYING THE CONNOTATIVE MEANINGS OF 'VIOLENCE'

The Word	Frequency
kill	4
slaughtered	2
attack	10
executed	2
blood	6
shot	4
destruction	1
confessed	6
killed	5
destroyed	5
crush	1
suicide	1
tortured	1
torn	2
tore	1
executions	2
slaughter	1
bloodshed	1

Table II displays the words that encode the associative meaning of the lexical item 'violence'. One observation is that all words included in the above table are ideologically-loaded. That is, they are carriers of specific meanings pertaining to specific theme, which is violence. The words in the table direct the readers to one meaning, that is, there is violence among discourse participants. All words connote the violent meanings of torture and suffering.

The theme of 'rebellion' can also be detected by the FDA as is shown in the following table.

TABLE III.   FREQUENCY DISTRIBUTION ANALYSIS OF THE WORDS CARRYING CONCEPTUAL AND ASSOCIATIVE MEANINGS OF 'REBELLION'

The Word	Frequency
rebellion	29
remove	1
expel	2
expulsion	8
uprising	1
striking	1
quarrel	2
quarrelling	1
dismissed	1
disobey	1
disobedience	1
rebelliousness	1
revolutionary	1
get rid of	2

Table III indicates that Orwell enriches the text with words indicating the meaning of rebellion, which is one of the main themes presented in Animal Farm. All lexis in the above table communicate the meaning of rebellion, both literally (e.g., rebellion, uprising, rebelliousness, revolutionary) and associatively (e.g., remove, get rid of, disobedience, dismissed, expel, striking). Further, it is not only high frequency words that communicate the theme of rebellion, but also low frequency words are carriers of the same theme.

A further theme presented in the novel can be also identified via the FDA and the number of occurrences of specific words. It is the theme of 'oppression' as is displayed in the following table.

TABLE IV.   FREQUENCY DISTRIBUTION ANALYSIS OF THE WORDS CARRYING CONCEPTUAL AND ASSOCIATIVE MEANINGS OF 'OPPRESSION'

The Word	Frequency
destruction	1
slavery	3
miserable	5
cruelty	2
traitor	3
laborious	3
criminal	1
execution	1
kill	4
destroy	1
execute	2
destroyer	1
dominate	3

Table IV clarifies the various words carrying the literal and connotative meanings of oppression. The semantic potentials of the list of words in the above table refer to the theme of oppression and domination. All words in this list carry meanings that indicate suffering, submission and dominance; such meanings are much more pertinent to the theme of oppression than to any other themes in the discourse of the novel.

C. FDA as Identifier of Indicative / Non-indicative Occurrences

This part of the analysis sheds light on one important idea beyond FDA adopted in this paper, that is, the significance of both high frequency and low frequency words in communicating the various pragmatic and stylistic purposes in discourse. Unlike the general assumption that only high frequency words are significant in delineating the various discursive aspects of texts, the paper shows that low frequency words have the same significance in producing and maintaining specific discourse meanings and themes. Consider the following Tables V and VI.

TABLE V.   INDICATIVE HIGH FREQUENCY WORDS IN ANIMAL FARM

The Word	Frequency
comrades	55
enemy	12
comrade	41
work	72
rebellion	29
man	21
all	174
no	102

TABLE VI.   INDICATIVE LOW FREQUENCY WORDS IN ANIMAL FARM

Word	Frequency
equal	8
friends	5
blood	6
equality	2
revolutionary	1
uprising	1
expulsion	8
remove	1
friendship	1

Tables V and VI demonstrate a number of words with high frequency occurrences (Table V), and other words with low frequency occurrences (Table VI). In both cases, the words contribute to the general interpretation of the text under investigation. For example, the words work, comrades and rebellion show high frequencies of 72, 55, and 29, respectively. The indication here is that the high frequency of occurrences pertaining to these words communicates various discourse meanings, such as equality, cooperation and friendship. These meanings, in turn, are indicators of the type of discourse as well as the theme intended to be conveyed pragmatically and/or stylistically by the writer. Likewise, the very low frequency of occurrences of words, such as equal, uprising, blood, remove and expulsion (Table VI) does not mean that these words are insignificant and, thus, do not contribute to the discourse meanings of the novel. Contradictorily, these words, however, low in frequency, are highly contributive to the identification of the thematic message of the novel at hand.

D. FDA as Lemmas Generator

A frequency distribution analysis can further be perceived as a generator of lemmatization. These lemmas are also indicators of the pragma-stylistic meanings targeted in the novel, as they refer to the various discursive and thematic purposes beyond the text in which they occur. Consider the following Tables VII and VIII.

TABLE VII.   LEMMATIZATION IN ANIMAL FARM

Destroy lemma		Execute lemma		Kill lemma		confess lemma	
word	F	word	F	word	F	word	F
destroy	1	executed	2	kill	4	confess	2
Destroyed	5	execution	1	kills	1	confessed	6
Destroyer	1	execution-s	2	killed	5	confessing	1
destroying	1			killing	1	confession	1
destruction	1					confessions	1

(F) means frequency

TABLE VIII.   LEMMATIZATION IN ANIMAL FARM

equality lemma		friend lemma		comrade lemma		rebellion lemma	
word	F	word	F	word	F	word	F
equal	8	friend	4	comrade	34	rebellion	29
equally	3	friends	2	comrades	46	rebellions	1
equality	2	friendly	3	comrade-ship	1	rebelling	1
		friend-ship	1			rebelliou-sness	1

(F) means frequency

The two tables display the lemmas of various indicative words in the discourse of the novel. These lemmas are indicators of the stylistic way of writing adopted by the writer. The selection and use of these words, together with their different lemmas is dexterously employed to direct the cognitive background of the reader towards specific meanings that serve to arrive at the pragmatic interpretation of texts. Lemmas are very indicative in indentifying the type of discourse as well as in classifying the different themes in texts.

V. DISCUSSION

The above analysis demonstrates the effectiveness of using and applying CATA represented by FDA to the linguistic analysis of texts. It is analytically clarified that the use of modern technology in the linguistic and textual analysis of texts, particularly literary ones contributes effectively to the interpretation of these texts. In light of this paper, the application of FDA proves useful and contributive in creating and deleneating the general interpretative atmosphere of the novel under investigation. This variable of CATA facilitates the process of decoding the various pragmatic and stylistic meanings pertaining to the text at hand. Themes such equality, violence, oppression and rebellion have computationally been decoded by means of FDA. This correlates with a number of previous studies, such as [28], [29] and [25], which emphasize the importance of applying modern technologies to the textual and linguistic analysis of texts. Crucially, computer software, when used in corpus linguistic, function to facilitate the process of linguistic analysis, as they help make texts more manageable analytically (Research question No. A: to what extent does a computer-based frequency analysis contribute to clarifying the pragma-stylistic meanings pertaining to particular words in the selected novel?).

Pragmatically, the application of CATA software in general and FDA in particular proves useful in deciphering the various pragmatic meanings pertaining to the discourse of Animal Farm. Pragmatic meanings are meant to the implied or the invisible meanings. This is conducted by using FDA as an indicator to mirror the intended meanings targeted by the writer. For example, FDA has shown certain pragmatic meanings relevant to the idea of totalitarianisn, oppression and submission, which represent the main meanings intended beyond the surface meanings of the linguistic expressions in the novel [30]. Deciphering the intended meanings of the work under investigation by means of FDA also serves to reflect the intention of the writer, which is further meant by the ideological point of view (Research question No. B: How does a frequency distribution analysis communicate specific pragmatic and stylistic meanings in the novel under investigation?).

Stylistically, the use of FDA demonstrates the way the writer employs certain lexis that communicate specific meanings beyond the style of his writing. The analysis shows two types of discourse: the discourse of equality and the discourse of inequality. Each type of discourse is featured by a number of stylistic devices manifested in the clever use of specific words that direct the interpretative wheel of the text towards the targeted type of discourse. Consequently, one conclusion can be drawn here, that is, CATA software can be used to determine the type of discourse in texts. This finding goes in conformity with some previous studies, such as [20], [23] and [31], who clarified that the application of CATA to the linguistic study of literary texts serves to identify the nature of discursivity in literary genres (Research question No. B: How does a frequency distribution analysis communicate specific pragmatic and stylistic meanings in the novel under investigation?).

Thematically, the application of FDA helps classify the different themes addressed in literary texts. Such thematic categorization is highly required, particularly in literary genres, for the very nature these texts have concerning the large data they contain. The thematic classification can also facilitate the process of teaching and learning literary courses. This reconciles with [32], who emphasizes the significant contribution computer software can present to EFL settings. The application of these software can save time and effort on the part of both teachers and students. It also tunes with [33], who argue that CATA software prove contributive to the acquisition of the different language skills, particulaly, reading, writing, as well the acquistion of vocabulary. The use of computer software serves to improve students' performance and competency (Research question No. C: How do high frequency words mirror the thematic and discursive nature of the analyzed text?).

Crucially, the huge technological development necessitates the integration of computer softwar not only in EFL settings, or in the linguistic study of literary texts, but also in the linguistic investigation of further types of texts, such as legal and religious texts. The application of the various CATA software produce credible, authentic and concise results [34]. Using computer software also opens new analytical and pedagogical insights that would be difficult to be identified if the analysis is conducted without the help of computer. This last point was accentuated by [35], who shed light on the different theoretical, analytical and pedagogical horizons computer software offer for researchers in the different fields of the academic and scientific research (Research question No. D: Does the use and application of a frequency distribution analysis facilitate an intelligible perception and understanding of various types of texts?).

VI. Conclusion

This paper presented a computer-based frequency analysis to decode the pragma-stylistic meanings in Orwell's Animal Farm. The paper demonstrated the significance of using and applying computer software in general and FDA in particular to the linguistic study of literary texts. These software function to save both time and effort, as well as provide results that are more credible, accurate and concise than those arrived at by traditional ways of linguistic analysis (i.e. without the work of computer). The analysis of the current paper clarified that FDA proves useful in (i) identifying the types of discourse in the novel at hand; (ii) categorizing the various themes in discourse; and (iii) highlighting the indicative and non-indicative occurrences that communicate different pragmatic and stylistic purposes in the novel under investigation. These three analytically-evidenced findings are computationally enabled by the application of FDA. The paper emphasizes the findings revealed by previous studies, by highlighting the significance and necessity of using the various computer software in the linguistic analysis of texts, particularly large data fictional texts. This is because these software facilitate the whole process of analysis, open new analytical horizons in the field, improve textual and contextual intelligibility pertaining to texts, provide fast, credible and concise results, and mirror the pragmatic and stylistic meanings communicated by writers.

Finally, for future research, this paper recommends further applications of the different computer software to the analysis of texts other than the literary ones. For example, to investigate the effectiveness of GBL (Game-Based Learning) on the performance of EFL (English as a Foreign Language) majors concerning vocabulary acquisition, or investigating the impact of CAT (Computer Assisted Translation) Trados Studio software on the teaching and learning translation. These recommended studies might reveal similar and/or different results than those approached in the current paper. Crucially, integrating computer software in the EFL settings contributes significantly to the teaching methods on the part of instructors, and to the learning outcomes on the part of students.

References

[1] G. Stockwell, Computer-assisted language learning: Diversity in research and practice. Cambridge: Cambridge University Press, 2018.

[2] A. F. Khafaga, "The perception of blackboard collaborate-based instruction by EFL majors/teachers amid COVID-19: A case study of Saudi universities". Journal of Language and Linguistic Studies, vol. 17, no. 2, pp. 1160-1173, 2021.

[3] D. Wiechmann, and S. Fuhs, "Concordancing software". Corpus Linguistics and Linguistic Theory, vol. 2, no. 2, pp. 107-127, 2006.

[4] D. Krieger, "Corpus linguistics: What it is and how it can be applied to teaching". The Internet TESL Journal, vol. IX, no. 3, pp. 123-141, 2003.

[5] K. Romeo, "A web-based listening methodology for studying relative clause acquisition". Computer Assisted Language Learning, vol. 21, no. 1, pp. 51-66, 2008.

[6] J. Flowerdew, "Concordancing as a tool in course design". System, vol. 21, no. 2, pp. 231-244, 1993.

[7] F. Yavus, "The use of concordancing programs in ELT". Procedia-Social and Behavioral Sciences, vol. 116, pp. 2312-2315, 2014.

[8] I. Pollach, "Taming textual data: The contribution of corpus linguistics to computer-aided text analysis". Organizational Research Methods, vol. 15, no. 2, 263–287, 2012.

[9] H. Bergqvist, "Swedish modal particles as markers of engagement: Evidence from distribution and frequency". Folia Linguistica, vol. 54, no. 2, pp. 469-496, 2020.

[10] A. F. Khafaga, and I. Shaalan, "Mobile learning perception in the context of COVID-19: An empirical studyof Saudi EFL majors". Asian EFL Journal, vol. 28, no. 1.3, pp.336-356, 2021.

[11] J. Reddington, F. Murtagh, and C. Douglas, "Computational properties of fiction writing and collaborative work". International Symposium on Intelligent Data Analysis, pp. 1-13, 2013.

[12] Q. Ma, "From monitoring users to controlling user actions: A new perspective on the user-centred approach to CALL". Computer Assisted Language Learning, vol. 20, no. 4, pp. 297-321, 2007.

[13] S. Hockey, A guide to computer applications in the humanities. London: The Johns Hopkins University Press, 1980.

[14] G. Kennedy, An introduction to corpus linguistics. London & New York: Longman, 1998.

[15] A. Barger, and K. Byrd, "Motivation and computer-based instructional design". Journal of Cross-Disciplinary Perspectives in Education, vol. 4, no. 1, pp. 1-9, 2011.

[16] K. Beatty, Teaching and researching computer-assisted language learning. Harlow: Longman Pearson, 2010.

[17] A. F. Khafaga, and I. Shaalan, "Using concordance to decode the ideological weight of lexis in learning narrative literature: A computational approach". International Journal of Advanced Computer Science and Applications, vol. 11, no. 4, pp. 246-252 , 2020.

[18] A. F. Khafaga, and I. Shaalan, "Pronouns and modality as ideology carriers in George Orwell's Animal Farm: A computer-aided critical discourse analysis". TESOL International Journal, vol. 16, no. 4.2, pp. 78-102, 2021.

[19] R. Dzekoe, "Computer-based multimodal composing activities, self-revision, and L2 acquisition through writing". Language Learning & Technology, vol. 21, no. 2, pp.73-95, 2017.

[20] A. F. Khafaga, Strategies of political persuasion in literary genres: A computational approach to critical discourse analysis. Germany: LAMBERT Publication, 2017.

[21] A. Thabet, "Applied computational linguistics: An approach to analysis and evaluation of EFL materials". Damietta Faculty of Education Journal, vol. 1, no. 13, pp. 7-39, 1990.

[22] J. Sinclair, Corpus, concordance collocation. Oxford: Oxford University Press, 1991.

[23] A. F. Khafaga, "Exploring ideologies of function words in George Orwell's Animal Farm". Pertanika Journal of Social Sciences and Humanities, vol. 29, no. 3, pp. 2089 -211, 2021.

[24] A. F. Khafaga, "A computational approach to explore the extremist ideologies of Daesh discourse". International Journal of Advanced Computer Science and Applications, vol. 11, no. 8, pp. 193-199, 2020.

[25] A. F. Khafaga, and B. Aldossari, The language of persuasion in courtroom discourse: A computer-aided text analysis". International Journal of Advanced Computer Science and Application, vol. 11, no. 7, pp. 332-340, 2021.

[26] A. F. Khafaga, and M. Aldawsari, "Ideological agency in Edward Bond's Lear". Applied Linguistics Research Journal, vol. 5, no. 2, pp. 11-23, 2021.

[27] A. F. Khafaga, and A. Alghawli, "The impact of CALL software on the performance of EFL students on the Saudi university context". International Journal of Advanced Computer Science and Application, vol. 12, no. 7, pp. 304-312, 2021.

[28] C. A. Chapelle, and J. Jamieson, Tips for teaching with CALL: Practical approaches to computer-assisted language learning. White Plains, NY: Pearson Education, 2008.

[29] L. Pedro, C. Barbosa, and C. Santos, "A critical review of mobile learning integration in formal educational contexts". International Journal of Educational Technology in Higher Education, vol. 15, no. 1, 2018.

[30] W. F. Bolton, The Language of 1984. Orwell's English and Ours. Basil Blackwell Publisher Limited, 1984.

[31] J. Flowerdew, "Concordancing as a tool in course design". System, vol. 21, no. 2, pp. 231-244, 1993.

[32] C. Pim, "Emerging technologies, emerging minds: Digital innovations within the primary sector," in: G. Motteram (Ed.), Innovations in learning technologies for English language teaching, London: British Council, pp. 17-42, 2013.

[33] J. Jarvis, and I. Pastuszka, "Electronic literacy reading skills and the challenges for English for academic purposes". CALL-EJ Online, vol. 10, no. 1, 2008.

[34] R. S. Pinner, "Teachers' attitudes to and motivations for using CALL in and around the language classroom". Procedia-Social and Behavioral Sciences, vol. 34, pp. 188-192, 2012.

[35] A. Oskoz, and I. Elola, "Digital stories: Overview". CALICO Journal, vol. 32, no. 2, pp. 155-173, 2016.

AUTHORS' PROFILE

Ayman Khafaga is an Associate Professor of Linguistics at the department of English, College of Science &Humanities, Prince Sattam Bin Abdulaziz University, Saudi Arabia. He is also an Associate Professor of Linguistics, Faculty of Arts & Humanities, Suez Canal University, Egypt. His research interests include computational linguistics, discourse studies, semantics, pragmatics and stylistics. ORCID: https://orcid.org/0000-0002-9819-2973.

Iman El-Nabawi Shaalan is an Associate Professor of Applied Linguistics at the College of Science and Humanities, Department of English , Prince Sattam bin Abdul-Aziz University, Saudi Arabia. She is also an Assistant Professor of TEFL at the College of Education, Al-Azhar University, Cairo, Egypt. Her research interests include language teaching methodologies, curriculum development and teaching methods, second and foreign language teaching and learning, and translation. ORCID: https://orcid.org/0000-0002-5411-7613.