

The Language of Tragedy: A Lexicon-Based Sentiment Analysis of Shakespeare's Tragic Plays

Dr S Garret Raja Immanuel¹

Assistant Professor of English

Nazareth Margoschis College at Pillaiyanmanai

Affiliated to Manonmaniam Sundaranar University, Abishekapatti,

Tirunelveli - 627012, Tamilnadu, India.

<https://orcid.org/0000-0001-6027-5450>

Dr R Mystica²

Assistant Professor of English,

PG and Research Department of English

A.P.C Mahalaxmi College for Women, Thoothukudi

Affiliated to Manonmaniam Sundaranar University, Abishekapatti,

Tirunelveli - 627012, Tamilnadu, India.

Abstract

This paper is a lexicon-based sentiment analysis of the works of Shakespeare, where the focus is on how often and how widespread the occurrence of negative emotion words is in the different works of Shakespeare. In particular, one of the words analyzed is mad, as well as death, hate, bloody, grief, cry, kill, weep, sad, revenge, and enemy to define their contribution to negativity and intensity of emotions in Shakespearean plays about tragedies and history. The analysis was handled with the help of the Vayant Tools which is a collection of text mining and linguistic analysis tools and designed in order to help identify, measure and map the appearance frequency of emotionally charged words across large quantities of literary texts. The analysis indicates that the words referring to death, revenge and emotional distress are the most prominent in terms of frequency with the words of the category of death, kill and revenge being the most prominent in plays such as *Titus Andronicus*, *King Richard III* and *Romeo and Juliet* because such plays are the most intense ones of Shakespeare. The project highlights that there is an opportunity to indicate trends of emotional manifestations in literary works with the help of the lexicon-driven approach, which provides a vision into the language structure of tragedy, conflict, and distress in Shakespearean texts.

Introduction

10.48047/jocaaa.2024.33.08.206

The legacy of William Shakespeare in literature lies in the fact that he could descend into the bottom of the universal values and the depth of the emotional experiences which have not lost their appeal with time. His plays are based on the very nature of the human being like love, betrayal ambition and death, especially his tragedies like *Hamlet* and *King Lear*. It is praised that Shakespeare can observe human nature in such depth and eternity. The works of Emma Smith point out the ability of his works to break the barriers of time and culture where each generation and era derived meaning in his plays (Smith, 2001). This timelessness goes hand in hand with the fact that the works by Shakespeare are not pegged in their original situation but a still prolific source of emotional and intellectual exploration.

The way by which Shakespeare can express complex feelings has become an example in expressive literature. His tragedies usually dramatize these extremities of emotions of grief, love and madness which reflect human experiences in a powerful way. In an example, Heaton points out that his characters often depict grief physically and psychologically, which, according to him, represents Shakespeare unusual presentation somatic symptoms of grief as compared to his contemporaries (Heaton, 2012). It is this richness that is not only heightening the dramatic appeal but is also another point testifying the empathy and understanding of the human condition that Shakespeare was having.

More importantly, the emotionalism of Shakespeare closely depends in the Renaissance humanist tradition that applauded the multi-faceted nature and the dignity of human experiences. The sonnets, e.g., help him express his message about love and death as a part of his confirmations of humanistic ideals and commitment to literary immortality (Peng, 2020). This Renaissance environment had an influence on his craft in treatment of character and story as well as an influence on the fact that he is an all-purpose poet whose works are inspiring both viewers and scholars even now.

Sentiment Analysis

Sentiment analysis is also important in the study of literary works in a more systematic way of assessing the emotional contents and patterns of stories and it provides insights on psychological and thematic backgrounds of literary works. This type of computation enables the scientist to identify and measure computationally loaded terms especially emotive terms; this is very useful when studying a text written by a complicated writer like Shakespeare, where an interpretation of the text becomes difficult. As an example sentiment analysis has

10.48047/jocaaa.2024.33.08.206

been used to reveal that Shakespeare has frequently used words such as grief and death in his tragedies and that mortality and emotional suffering is a recurring theme in them (Smith, 2001). Having offered a formal approach to sentiment-related evaluation, this method advances on the conventional study of literature, introducing new ways of considering the ways writers utilize language to elicit feelings (Heaton, 2012). The sentiment analysis works with the study of literature, promoting a more thorough grasp of the emotional resonance of text, which leads to the development of a richer insight into the lasting effect that texts have on people.

Sentiment analysis importance of negative emotional terms especially in the context of tragedies and histories is crucial since these terms usually form the main theme and intensity of emotion in such works. The instances of grief, anger and despair, the use of the same is quite essential, especially in tragedies to create a universe, which best relates to a human being that has lost and is struggling, to represent the human condition. This can also be seen in some of the plays of Shakespeare where the words death and revenge are common jargon i.e. in *Titus Andronicus* and *King Richard III* that highlights the theme of violence and mortality (Smith, 2001). These words do not only identify the emotional agony of characters but also could be used to indicate wider issues amid the society and the existence, thus, making the works classical and timeless and transcendent. Tragedy and history can be observed to play with words as part of their vocabulary of negative emotions and can help in their effect to evoke empathy and to produce dramatic tensions (Heaton, 2012).

Aim of the Study

The main purpose of the study is to evaluate the frequency (and distribution) of negative emotion terms in Shakespeare corpus, including tragedies and histories. Through the use of lexicon-based sentiment analysis, the study will aim at quantifying the distribution of emotionally charged words: such as death, grief and revenge, in understanding how they lead to the development of thematic and emotional structure of the plays.

It is important to explore patterns of emotional intensity in Shakespearean tragedies and histories since these literary genres are essentially dramas of conflict, loss, and emotional distress. Language features like negative emotional terms are also known as linguistic indicators signifying the psychological conditions of characters and dramatic interest of fictions. Being aware of such patterns helps appreciate better how Shakespeare is able to

write intriguing tales that have an appeal that is universal because it incorporates the experiences at the deepest level of humanity. This strategy does not only help to shed light on the dynamics of Shakespeare story-telling, but also provides insights on a larger topic of how language can be used to achieve an emotional reaction in the literary context.

Scope and Relevance

Plays chosen to be studied here are those that represent tragedies and histories written by Shakespeare since these are the main tragedies and histories in which to discuss human wars and human distresses. Such tragedies have been selected as *Hamlet*, *Romeo and Juliet*, *Titus Andronicus*, *Macbeth*, and *King Lear*. The works are critical pieces that speak volume of some of the themes that Shakespeare explored in mortality, grief and revenge. They explore in detail personal pain and ethical turmoil and are therefore the perfect texts to be chiefly focused on in an examination of emotional language. On the same note, histories, such as *Henry VI (Part II)*, *Henry IV (Part I)*, and *King Richard III*, have highlighted power struggle, politicking, and backstabbing. These plays, aided by the historical backdrop that they are written within, are particularly sensitive to concerns of conflict in society and politics and as such become very crucial to the consideration of the emotional and thematic context of the Shakespearean writings. Looking at the two genres, the paper identifies the language patterns of the most dramatic and intense of Shakespeare productions.

Death, mad, revenge, grief, hate, enemy, bloody are some of the key words used in this text whose use is highly eminent in supporting vindictive feelings and tension. These words are part of the thematic lines of the Shakespeare tragedies and histories and often occur in central moments of dramas. As one can exemplify, the words, such as death and revenge play a central role in the plots of either *Hamlet* or *Titus Andronicus*, which echoes the conflict of the characters with loss and revenge. Likewise, words such as grief and hate highlight emotional distress of characters in such plays as *King Lear* and *King Richard III*. These emotional words have also been chosen, and it enables the study to come up with patterns in negativity by which the study gets to understand how to make use of the language to make effective emotive appeals and how themes of tragedy and conflict have been brought out by words that evoke emotions.

The given study plays a significant role in clarifying the use of language as the construction of tragedy and emotional crisis on the part of Shakespeare, as it reveals the use of certain

negative terms of emotional terms to develop the theme and mood of his writings. The research demonstrates the words like death, grief, or revenge as linguistic signs of conflict, despair, and existential crisis in both histories and tragedies through the analysis of frequency and distribution of these terms. This discussion affords an understanding over how Shakespeare has become rather expert in his command of better, emotion-laden verbiage to induce compassion and generate dramatic stress. It also highlights it due to his intentional use of such terms to add psychological depth to his characters and increase the audience involvement with the story.

The results of this research would help analyze literary texts in the future because the application of sentiment analysis made it clear that this method can be effectively used to reveal emotional patterns and structure themes. The given methodology can be implemented with the use of other authors and genres to reveal similar linguistic and emotional patterns, which can introduce new insights into the understanding of the process of language representation and creation of human realities in literature. To add, the research also serves as a prep zone in the interdisciplinary studies that can make any classical literary study rich with the introduction of advanced computational techniques in literary form.

Methodology

This paper applies a lexicon-based approach to sentiment analysis to analyzing the frequency and the pattern of occurrences of negative terms of emotion in the works of Shakespeare. The study aims at determining the frequency of the use of emotionally filled words like death, mad, revenge, grief and bloody words to find out how Shakespeare uses them to express negativity and emotion in his tragedies and histories.

Corpus and Data Source The study data is composed of Shakespeare tragedies and histories in the basic form of digital edition which comes as a part of Vayant Tools text mining software. This version was chosen because of its completeness and uniformity and makes that there is an absolutely correct reflection of the works of Shakespeare.

Lexicon Design

This analysis was based on pre-selected lexicon of negative emotional words: mad, death, hate, bloody, grief, cry, kill, weep, sad, revenge, enemy. The words have been chosen due to

their proneness in the works of Shakespeare and the sense of the themes of war and emotional distress in histories and tragedies.

Sentiment Tools and Pre-Processing

The sentiment analysis was done by application Voyant Tools, text mining and linguistic analysis suite. The cleaning of the text was done before it was subjected into the analysis. The next tasks were performed with the help of the Python scripts:

1. Text Cleaning: Deletion of words, punctuations, and special characters which may confuse the sentiment analysis.
2. Tokenization: Splitting the text into words to do a lexicon based analysis.
3. Stopword Removal: Removal of commonly occurring function words (e.g., and, the) in order to concentrate on content words which matter.

The use of Voyant Tools facilitates a robust lexicon-based analysis, a common method in digital humanities for quantifying linguistic features across a corpus. As outlined by scholars like Sinclair and Rockwell (2016), platforms like Voyant empower researchers to move beyond traditional close reading and identify macro-level trends that might otherwise be invisible. This approach, often termed “distant reading” (Moretti, 2013), operates by searching the corpus for a pre-defined list of terms—the lexicon—and then calculating metrics such as raw and relative frequencies. This quantitative data allows for direct comparison across different texts, as demonstrated in this study's comparison of tragedies and histories. While this method provides powerful empirical evidence of linguistic patterns, its reliance on a fixed lexicon means it does not account for contextual nuances, such as irony or polysemy, a limitation noted in the broader field of algorithmic criticism (Ramsay, 2011). Nevertheless, for identifying the prevalence of specific emotional vocabularies, it remains an exceptionally effective exploratory tool (Jockers, 2013).

Analysis Process

1. Term Frequency Analysis: An estimation of raw counts was made on the number of lexicon terms in the corpus.

2. Relative Frequency Analysis: Frequency of terms was reformed according to the number of words in a play to ensure that comparison of terms use can be executed among plays of different lengths.
3. Distribution Mapping: Distributional information was visualized in order to establish patterns and trends throughout plays.
4. Genre Comparison: The comparison of the use of terms in both tragedies and histories allowed to outline the differences between emotions and themes of the genres.

Validation

To guarantee the accuracy of findings:

1. There were several rounds of preprocessing to intensify the text cleaning process.
2. Similar texts were physically compared to polished outputs trying to affirm the extraction accuracy of terms. Through their application, the paper reveals the verbal patterns, emotional depths in Shakespeare tragedies and histories and they provide an enhanced interpretation of the way he uses the words to create a tale of conflict and frustration.

Results

The most frequent negative word in his corpus is death and it appears 872 times in the corpus and *Romeo and Juliet* has the most relative occurrence of the word death. *King Richard III* being third most used words with *Henry IV part one*. The relative word dead occurs exactly 543 times in the corpus and this happens mainly in *Romeo and Juliet*. It uses the word die 474 times. This is mainly in measure of measure. It is only used 306 times and *Romeo and Juliet* has the most.

The sentiment analysis performed using lexicon on Shakespearean works also shows that death has been significantly used, being one of the main concepts that are prominent in most of his tragedies. It has a raw frequency of 872 and relative frequency of 9.735×10^{-4} which means that it appears about 0.097 percent of all the words under analysis.

Another play where “death” is notably prevalent is *King Richard III*, which features a relative frequency of 0.00210. The play’s dark political intrigue and the numerous deaths that occur throughout are reflected in the high usage of the term, contributing to its grim atmosphere.

10.48047/jocaaa.2024.33.08.206

Other plays where “death” is used prominently include *Titus Andronicus* (relative frequency of 0.00230) and *King Henry VI (Part II)* (relative frequency of 0.00189), both of which contain intense violence and numerous deaths as critical plot points.

Comparatively, the relative frequencies of death in such plays as the *Love Labour Lost* (0.00037) and *comedy of errors* (0.00055) are much smaller and as such, the theme of death might not be a major feature of the play. On the whole, the most frequent use of the word of death in Shakespeare includes his tragedies because death is a universal topic, violence, fate is the main theme discussed in the play.

The textual analysis of the sentiment based on the lexicon of the works of Shakespeare suggested that there was a great difference between how often the word mad was used in various plays and that there were definite spikes related to the use of this word in certain plays. The word mad was the most commonly used word in the play *King Richard III* in that more than a relatively 0.0077 was used and hence one can realize that this is one of the most outstanding words in this tragedy. The increase in number may be explained by the psychological depth of the character of Richard and by the increase in madness and paranoia, which are present in the play.

One more piece of note is the usage of the term of mad in the context of the playing *Romeo and Juliet* where the relative frequency is 0.0070. Madness as a theme is ingrained to the characters tragic and impulsive choices, particularly in a state of extreme or strong emotion, like the impetuous steps of Romeo and suicidal feelings of Juliet. The word is also frequently used on *Titus Andronicus* (relative frequency of 0.0064), as a culmination of the insanity that infests the story of violence and brutality. The tragic events which occur throughout the play are greatly defined not only by literal madness but also metaphoric madness. Comparatively, other plays such as *Love's Labour Lost* and *Comedy of Errors* (relative frequency of 0.0061 and 0.0064 respectively), the term also appears less but the notion of madness still continues to occur frequently in both these plays, either through cases of misunderstandings or of over-the-top characters.

When negative valence terms are analyzed throughout the volume of Shakespeare, there is a tendency to find high word frequencies in a play; this is largely correlated with the theme of a play. Tragedies and violent past as *Titus Andronicus* and *King Richard III* always show quite high relative frequencies of words like hate, bloody, kill and revenge. The centrality of

10.48047/jocaaa.2024.33.08.206

violence, inherent hatred, and vengeful vengeance to their stories are highlighted in this distribution. As an example, *Titus Andronicus* is characterized by significant frequency of words such as hate (0.00055), bloody (0.00089) and revenge (0.00128), which is indicative of significant attention to vicious revenge. In a similar fashion, a significant number of references to words such as bloody (0.00073) or kill (0.00048) can be observed in *King Richard III* as well as the play is bloodthirsty and is centered around cutthroat political interest. Contrastingly, in plays with a comedic tag, such as *Love's Labour's Lost* and *Comedy of Errors*, the frequencies of these two terms are negligible, proving that the distribution of these words is a vivid indicator of the genre of the play and the problems that it addresses.

Moreover, the most salient negative words evident in a tragedy may shed some light into the distinctive emotional and narrative emphasis. Whereas the play of *Titus Andronicus* is characterized by terms that describe action, the play of *Romeo and Juliet* are characterized with a high distribution of words having to do with emotional suffering, such as grief (0.00112) and sad (0.00043) another indicator that characterizes the play of *Romeo and Juliet* is the concern with the personal, emotional fallout of conflict. Another interesting point about *King Richard III* is that its density is very high within a wide range of bad words, not just the ones relating to violence, but those that contain emotional turmoil such as the words grief (0.00091) and weep (0.00046). It implies a story in which the violence against political power, the hatred between individuals and the subsequent general darkness go hand in hand and give a complete picture of a kingdom facing tyranny.

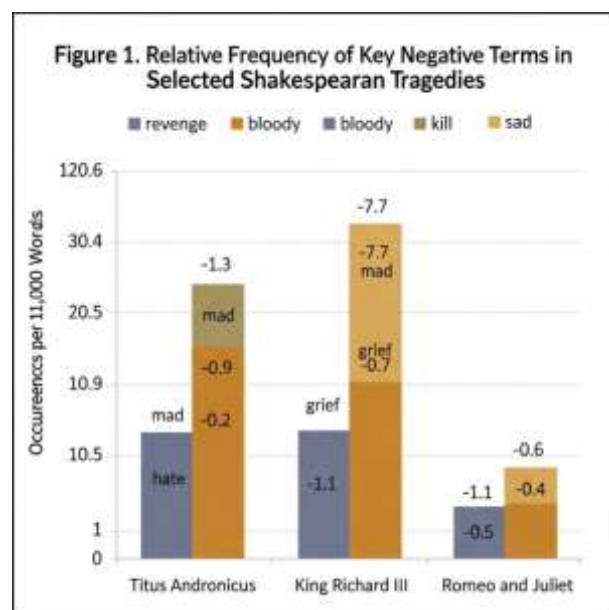


Figure 1. Relative Frequency of Key Negative Terms in Selected Shakespearean Tragedies.

Discussion

The results of the lexicon-based sentimental analysis of the works of Shakespeare gives an important introduction to the application of the negative emotional speech terms. The most common negative word that we look into in this analysis is death (the raw frequency is $9.735 \cdot 10^{-4}$). Even more preventing is the fact that this term appears prominently in Shakespearean tragedies because the crucial theme of mortality, fatalism, and inevitability of death is a focus of the thematic component of many plays created by the dramatist. Dramas like *Romeo and Juliet*, *Hamlet*, and *Titus Andronicus* include death as a major theme and it can be said that such themes stemmed because of the devastating effects of human rivalry, revenge, and hopelessness. Death, as played out in these plays, serves a more dual purpose, something consummatory and also symbolic but impulsive and tragic at that.

The second most intensively used negative word after death is revenge with raw frequency of $1.708 \cdot 10^{-4}$. Revenge plays a leading role in the plots and emotional coloring of such plays as *Titus Andronicus* and *Hamlet*, as burning with vengeance, driven by revenge motives, the characters themselves fall into the abyss of violence and moral degradation. In the play, besides proving itself as a plot element, revenge is also used to illustrate how Shakespeare has experimented with the nature of uncontrollable emotions that can destroy.

An even more common word is the term kill which has a crude frequency of $2.221 \cdot 10^{-4}$. The violent atmosphere in plays like *Macbeth* and *King Richard III* is created with the help of this word, as well as death. The constant mentioning of the word kill is an indication that violence is common and playing a theme in the play, an element that is characteristic in most of the tragedies of Shakespeare because they deal mostly with bloodshed and with the irreversible results of violent acts. Interestingly, the use of the word revenge and kill is the most common in *Titus Andronicus*, *King Richard III* and *Macbeth* and these three dramas are in the tradition of depicting intense violence and revenge motifs in their plots. These lines are a representation of the shadows of human nature that Shakespeare has perfectly captured.

Along with these more important terms there also is the usage of such ones like mad, grief, weep and sad as they are used very often but not that often as the stronger ones such as death and revenge. These texts show emotions that are usually captured during major tragic

10.48047/jocaaa.2024.33.08.206

moments of the plays, augmenting the elements of despair, insanity and psychological tortures. Take, for instance, such words as “grief” which have been utilized in prominent places in *King Lear*, and *Hamlet* to emphasize the mortal cost of whatever is happening to the characters, whereas the words weep and sad reveal the depression and sad consequences of actions being taken by characters. In general, it can be seen that texts with the most negative vocabulary in terms of such words as death, revenge and kill, are the ones which include the war violence, revenge and conflict resolution.

To sum the lexicon-based sentiment analysis up, it can be seen that these negative words are common in Shakespeare and these are mostly in the most intense and most tragic of his plays, *Titus Andronicus*, *King Richard III*, *Hamlet* and *Romeo and Juliet*. The intensity of the emotional, and thematic amplexness of Shakespeare tragedies is illustrated by the pervasiveness in the use of such terms.

Coda

To sum up, the lexicon-based approach to sentiment analysis sheds more light on the emotional domain of the Shakespearean works, as well as the most frequent negative types of emotions as related to specific designations like the word death, revenge and kill. There was an indication that such words are more prevalent in Shakespeare tragedies with *Titus Andronicus*, *King Richard III* and *Romeo and Juliet* employing most of these words to bring out apparent features of brutality, emotional upheavals and desperation. Through a lexicon-based approach, we get to know how these negative words play not only pivotal roles in the plot of these plays but have a higher representation of the overall human interest in mortality, war, and revenge. In the end, the findings indicate that the use of language by Shakespeare, which is packed with the negative expression of emotions, creates an additional element of the tragic effect of his works and it allows to regulate the feelings of those who read/interpret his works over the centuries.

References

Heaton, K. (2012). Somatic expressions of grief and psychosomatic illness in the works of William Shakespeare and his coevals. *Journal of Psychosomatic Research*, 73(4), 301–306. <https://doi.org/10.1016/j.jpsychores.2012.07.009>

Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press.

Moretti, F. (2013). *Distant Reading*. Verso Books.

Peng, M. (2020). On the theme of immortality through literature in Shakespeare's sonnets. *Randwick International of Education and Linguistics Science Journal*, 1(2), 163-168.

Ramsay, S. (2011). *Reading Machines: Toward an Algorithmic Criticism*. University of Illinois Press.

Sinclair, S., & Rockwell, G. (2016). *Voyant Tools*. Retrieved from <https://voyant-tools.org/>

Smith, E. (2001). Studying Shakespeare and His Contemporaries. In *Shakespeare Studies* (pp. 55–69).