

# Analysis of Antithetical Elements in English Literary Passages Using Stochastic Models

Chenjie Zeng

School of Humanities and Social Science  
The Chinese University of Hong Kong, Shenzhen  
Shenzhen, China  
E-mail: chenjiezeng@link.cuhk.edu.cn

Clement H. C. Leung

School of Science and Engineering &  
Guangdong Provincial Key Laboratory of Future Networks  
of Intelligence  
The Chinese University of Hong Kong, Shenzhen  
Shenzhen, China  
E-mail: clementleung@cuhk.edu.cn

**Abstract**—Thesis and antithesis are rhetorical figures often employed by well-known authors in English literary passages. Antithesis refers to the juxtaposition of contrasting words or ideas, which often, although not always, appears in the form of parallel structures. Such rhetorical figure is used to contrast opposing ideas, creating a sense of tension and urgency, as well as heightening the emotional impact of the speech. Contrasting the antithesis is the thesis, which refers to a statement, assertion, or tenet. A thesis also is a proposition laid down or stated, especially as a theme to be discussed and proved, or to be maintained against attack. It is observed that the points of occurrence of thesis and antithesis are random, yet they occur with remarkable regularity alternating each other. In this study, we represent the lengths of thesis and antithesis by using an alternating renewal stochastic process. We find that the underlying parameters of the alternating renewal process are able to usefully characterize the writing style of individual authors and help us to quantify and understand their linguistic technique and intention. Using the present method in conjunction with other techniques, such as sentiment analysis, and word embedding models, it is possible to gain a deeper understanding of the literature and its underlying themes and structures.

**Keywords**—stochastic alternating renewal process; antithesis; literary passages; thesis; renewal function.

## I. INTRODUCTION

Mathematical models provide a useful analytic mechanism for the quantitative study of the characteristics of literary work. For example, Markov models have been used extensively in studying English literature and language. One notable application is in the area of computational stylometry, which is a field that uses statistical methods to identify the authorship of anonymous or disputed literary texts. The use of Markov models in literature and language analysis has been a fruitful area of research, providing new insights into the structure and patterns of language use in literature, as well as enabling new methods for identifying core features of individual writers and texts.

For example, in Shakespeare's play *Henry V*, Henry's speech to his men before the Battle of Agincourt (Act 4, Scene 3) is a masterful example of rhetorical and dramatic effectiveness. The speech has its historical and literary value

and importance. On the one hand, the Battle of Agincourt is a triumph for the English in the Hundred Years' War. The battle occurred on Saint Crispin's Day, October 25th, 1415. Despite being outnumbered by the French, the English emerged victorious, and this unexpected win had a significant impact on English morale and reputation. It also dealt a severe blow to France and began a new phase of English superiority that would last for 14 years. On the other hand, Henry's speech is filled with rhetorical figures, such as antithesis, rhetorical questioning, allusion, parallelism, and so on. These stylistic devices create a stirring call to arms that inspires the English soldiers to fight with courage and conviction.

Therefore, the stylistic and rhetorical figures serve the theme of Henry's speech to his men before the Battle of Agincourt in Shakespeare's play *Henry V*. The theme is related to the power of leadership and inspiration in the face of adversity, and the responsibility of leaders to motivate and uplift their followers. The speech also shows personal responsibility, and the ability to rise to significant challenges. In this speech, Henry seeks to inspire his soldiers to fight and win against overwhelming odds, despite their fears and doubts. He does this by appealing to their patriotism, sense of honor and duty, and faith in God. He also shows them he is with them, leading the charge into battle and sharing their risks and struggles.

Through this speech, Shakespeare portrays the idea that leadership is not just about commanding others but inspiring and motivating them to be their best selves. It is about connecting with people on a deeper emotional level and giving them a sense of purpose and direction, even in the face of great danger and uncertainty. Henry's speech also highlights the importance of personal responsibility and accountability for leaders and those they command. By taking personal responsibility for his soldiers' well-being and showing that he is willing to share in their burdens and take on the risks of battle, Henry earns their respect and loyalty, and inspires them to fight with all their might.

In Section II, previous and relevant studies in the vast field of English literature and computer science will be introduced, so to portray a landscape of the related work. In Section III, we define the notation and function of "thesis" and "antithesis" as rhetorical devices in literary passages, including Shakespeare's *Sonnet* and *Henry V*, where King

Henry delivered a speech to his men before the Battle of Agincourt. More importantly, in this section, we explain how literary rhetoric is associated with mathematic and stochastic models, which bridges the gap between the disciplines of English Literature and Computer Science. Later in Section IV, as the nature of this interdisciplinary paper, we model the rhetoric thesis and antithesis by applying alternating renewal processes. Finally, in Section V, we test the model by conducting experiments. We analyze two literary passages in vastly different genres, including Leo Tolstoy's novel "War and Peace" and Barack Obama's 2008 presidential speech "Yes We Can".

## II. RELATED WORKS

Antithesis, known as confrontation or contradiction, refers to the juxtaposition of opposing thoughts, concepts, meanings, and images that are logically comparable [14]. It is a literary device characterized by the presentation of contrasting words or meanings. Many literary scholars have studied the function of antithesis. For example, in their work, Raximovna and Mirusmanovna [14] elucidate the practical characteristics of antithesis within literary texts, employing examples from both English and Uzbek fiction for analysis. It is commonly employed in literary texts, particularly in works of fiction, and holds significant value for in-depth examination [14].

Moreover, mathematical models have been used extensively in studying English literature and language. To be more specific, researchers have used Markov models to analyze the writings of Shakespeare, examining the patterns of word choice and syntax in his plays and sonnets to identify his distinctive writing style. By using Markov models to identify the probability of certain words or phrases occurring in a particular order, researchers have been able to detect subtle features of Shakespeare's writings [1]–[3]. Markov models have also been applied to study narrative structure in English literature. Researchers have examined the patterns of plot elements and character development in novels and other works, using Markov models to identify the most common transitions between different states in the narrative. This has allowed them to identify the key events and turning points in a story, as well as to analyze how different characters influence the course of the plot. For instance, Markov models are commonly used for statistical learning applications to capture the sequential patterns of data over time. While Hidden Markov models are widely researched, Devesh [7] explores an approach inspired by symbolic dynamics. This approach involves two main steps for successfully representing time-series data in a discrete space: first, continuous attributes are discretized, and second, the size of the temporal memory of this discretized sequence is estimated. Both steps are crucial for an accurate and concise representation of time-series data. The first step, discretization, determines the information content of the resulting sequence. The second step, memory estimation, is essential for extracting predictive patterns in the discretized data. The effectiveness of using a discrete Markov process for signal representation is determined by these two steps.

Another area of research has focused on using Markov models in natural language processing, which is the field of

computer science that involves using computers to process and analyze human language. Markov models have been used to analyze the structure and syntax of sentences, as well as to identify patterns of co-occurrence between different words or phrases in a text. For example, Eder et al. [6] discuss R in the context of Stylometry with R, used for analyzing writing styles in stylometry. Stylometry is a field that studies writing styles quantitatively, such as authorship verification, which can be useful in forensic contexts and historical research. The paper presents the potential applications of stylometry for computational text analysis, using several example case studies from English and French literature. The package is particularly effective in exploratory statistical analysis of texts, especially with regard to authorial writing style. The package has an appealing graphical user interface for novices without programming skills, such as those in the Digital Humanities. Experienced users can benefit from the package's standard pipelines for text processing and different similarity metrics.

## III. THESIS AND ANTITHESIS IN LITERARY PASSAGES

One of the key rhetorical figures used in the speech is the antithesis. Antithesis refers to the juxtaposition of contrasting words or ideas, which often, although not always, in parallel structure. For example, Shakespeare's Sonnet says, 'Before, a joy proposed; behind, a dream' [16]. This figure is used to contrast opposing ideas and create a sense of tension in "Henry V". The contrast of antithesis is thesis. According to Oxford English Dictionary [11], a "thesis" refers to a statement, assertion, or tenet. Thesis also is a proposition laid down or stated, especially as a theme to be discussed and proved, or to be maintained against attack. Thesis, in *Logic*, is sometimes as distinct from hypothesis, while, in *Rhetoric*, it is contrasted with antithesis. The formal connection between these figures demonstrates that despite their discrete, distinct, or potentially conflicting natures, they have the capacity and necessity to contribute to the larger entity of creation or society, existing alongside each other in a meaningful manner [15]. For example, when Henry says (in stanza 1, Act 4 Scene 3, *Henry V*),

		X	2	3	4	
If we are marked to <b>die</b> ,						we are enough
5	6	7	8	9	10 11 12	-X
To do our country loss; and if to <b>live</b> ,						
		X	2	3		-X
The <b>fewer</b> men, the <b>greater</b> share of honor						

Henry highlights the high stakes of the battle and contrasts the potential for death with the potential for glory. The highlights and contrasts create a sense of urgency and heighten the emotional impact of the speech. Shakespeare explores the concept of patriotism as a thematic element juxtaposed with profoundly unsettling notions, striving for a delicate equilibrium [12]. We use "X" to represent the first occurrence of rhetoric "die/fewer," and use "-X" to represent the follow-up antithesis "live/greater". Given the condition that an "X" appears in a poetic line, it is expected to have an

“-X” sooner or later. In the above lines that contain antithesis in Henry’s speech, when we come across “X,” such as “die” in the third line, we expect to have an antithesis “-X” like “live” later. Likewise, when we have the word “fewer”, which is an “X”. We will have its antonym “greater”, which can be symbolized as “-X”.

In order to analyze this phenomenon, let us count “X” and “-X” as 1, and then we can determine the exact length between them. The numbering starts from “X” and ends just before “-X” occurs. For instance, the beginning of the first “X” is “die”, which is number 1. The length in terms of the number of words between “X” (die) and its antithesis “-X” (live) is thus 12 words. The same length-counting goes on when we find the next “X” (fewer) and “-X” (greater). And the length between “fewer” and “greater” is 3 words. Given a simple calculation, we will have the average length  $(12+3)/2=7.5$  words in the first stanza of Henry’s speech to his men before the Battle of Agincourt.

Another example would be in Stanza 7 of Henry’s speech to his men before the Battle of Agincourt in *Henry V*:

X

Old men forget: yet all shall be **forgot**,  
 2 3 4 5 6  
 But he’ll **remember** with advantages  
 7 8 9 10 11 12  
 What feats he did that day.  
 13 14 15 16  
 Then shall our names,  
 17 18 19 20 21 22 23  
 Familiar in their mouths as household words,  
 24 25 26  
 Harry the king,  
 27 28 29  
 Bedford and Exeter,  
 30 31 32  
 Warwick and Talbot,  
 33 34 35  
 Salisbury and Gloucester,  
 36 37 38 39 40 41 -X  
 Be in their flowing cups freshly **remember’d**.

In stanza 7, the length between the first occurrence of the thesis X “forgot” in the first line and “-X” (remember’d) in the tenth line is 41 words. Provided that there is no other prominent antithesis in Henry’s speech to his men before the Battle of Agincourt, we can have an average length of the two stanzas:  $(7.5+41)/2=24.25$  words.

Furthermore, the same pattern of “X” and a follow-up “-X” can also be seen in rhetorical questions. The rhetorical question is usually defined as any question asked for a purpose other than to obtain the information the question asks. For example (at the beginning of stanza 1),

X

What’s he that wishes so?  
 -X X -X  
 My cousin, Westmoreland? No, my fair cousin;

Literarily speaking, “What’s he that wishes so?” and “Westmoreland?” are likely to be statements regarding one’s opinion of the person addressed rather than a genuine request to know. Similarly, when someone responds to the question by saying, “My cousin” is more likely to be an expression of feeling than a realistic request for information. The technical term for rhetorical questions, in general, such as “No, my fair cousin”, is erotema, which is a rhetorical question to affirm or deny a point strongly by asking it as a question.

If we call the whole sentence “What’s he that wishes so?” as “X,” then we are expecting an answer like “My cousin”, which is an “-X”. The same pattern goes in the second pair of “questioning” and “answering” followed up. The second “questioning” is “Westmoreland?”, which is a “X”, and the second “answering” is “No, my fair cousin”, which is another “-X”.

Furthermore, allusion is also an effective rhetorical device used in Henry’s speech. An allusion is a literary technique that makes a short reference to a well-known thing that the audience will likely know. This technique enables writers and speakers to convey a lot of meaning and importance in a concise manner. When Henry references the feast of Crispian, he draws on historical and religious symbolism to motivate his troops. The allusion serves to create a sense of continuity between the present moment and the past, and to elevate the importance of the battle in the minds of the soldiers. Nevertheless, the effectiveness of allusions depends on readers’ comprehension and recognition of them, and their correct interpretation of the associated significance. If an allusion is perplexing or misconstrued, it can reduce its effectiveness by perplexing the reader.

While comprehending the comedic aspect may not be entirely achievable, it is evident that a significant source of amusement for the audience lies in encountering familiar and unpretentious traditions presented in a manner typically associated with religious or moral contexts, thereby alleviating the need for restraint [13]. For example, Crispin and Crispinian, are twin brothers born to a noble Roman family, having been beheaded during Diocletian’s reign. They were executed for their religious beliefs on October 25th, probably in the year of 285 or 286. While fleeing religious persecution, the brothers worked diligently at their cobbler business in secret during night-time hours. By utilizing their trade, the brothers were able to support themselves and aid those in need. Their success led to Rictus Varus, governor of Belgic Gaul, becoming hostile towards them, resulting in the brothers being tortured and heaved into the river with millstones tied to their necks. Despite surviving this ordeal, they were killed by the emperor in 286. Later, October 25th becomes the feast day of Saints Crispin and Crispinian. Given the above historical background, it is expected that the twin brother “Crispin and Crispinian” should always appear in pairs. If we make “Crispin” as “X” and “Crispinian” as “-X”, we can see a similar and repetitive pattern of “X” and “-X” in the speech. For instance,

This story shall the good man teach his son;  
 -X X  
 And **Crispin Crispian** shall ne’er go by,

Similarly, the repetitive pattern is shown as follows as well.

X                      -X  
**Bedford and Exeter,**  
X                      -X  
**Warwick and Talbot,**  
X                      -X  
**Salisbury and Gloucester,**

Bedford and Exeter, Warwick and Talbot, and Salisbury and Gloucester are all pairs of names of places in England or surnames of people. Specifically, in Shakespeare's "Henry V", these names refer to specific people. Namely, the Duke of Gloucester and Duke of Bedford are brothers to the King; the Duke of Exeter is uncle to the King. And the following characters should be officers in King Henry's army, including Earl of Salisbury, Earl of Westmoreland, and Earl of Warwick. They are important figures and contrasts in the events leading up to and during the Battle of Agincourt. The next section develops a general stochastic model for these situations.

#### IV. MODELLING THESIS AND ANTITHESIS USING ALTERNATING RENEWAL PROCESSES

The alternating recurrence of Thesis and Antithesis can be modelled as an alternating renewal stochastic process [8]. Upon the occurrence of the thesis, a certain length of words ( $X_i$ ) elapsed before the occurrence of the antithesis; for such words, we shall refer to them as being under the dominance of the thesis. Likewise, upon the occurrence of the antithesis, a certain length of words ( $Y_i$ ) also elapsed before the occurrence of the thesis; for such words, we shall similarly refer to them as being under the dominance of the antithesis. By focusing on such rhetorical devices, we have the following model of a passage.

$$S = X_1 + Y_1 + X_2 + Y_2 + \cdots + \cdots X_N + Y_N \\ = \sum_{k=1}^N X_k + \sum_{k=1}^N Y_k ,$$

where  $\{X_i\}$  are independent, identically distributed positive random variables representing the length of words under the dominance of the thesis pending resolution by the antithesis, and  $\{Y_i\}$  are also independent, identically distributed positive random variables representing the length of words under the dominance of the antithesis awaiting the next occurrence of a new thesis or until the end of the passage is reached. In general,  $\{X_i\}$  and  $\{Y_i\}$  have different distributional properties. We shall refer to  $X_i$  as the thesis length and refer to  $Y_i$  as the antithesis length. The lengths of thesis and antithesis will form an important basic characterization of a piece of literary work.

Certain authors prefer to deploy a relatively lengthy  $X_i$  so that greater tension can be built up until its eventual resolution; other authors may choose to adopt a medium or short length for  $X_i$  to bring about a sharper impact. Thus, the  $\{X_i\}$  and  $\{Y_i\}$  often provide a useful mechanism to characterize the stylistics of different writers. For an arbitrary word under the dominance of a thesis, it can therefore be

classified as either positive, which corresponds to a continuation of the thesis dominance pending resolution or negative, which corresponds to a resolution that switches from a thesis dominance to an antithesis. Under the dominance of a thesis, the propensity of a word being positive is represented by the probability  $p$ , while the probability that it is negative is represented by the probability  $q$ , where  $p + q = 1$ . On the other hand, for an arbitrary word under the dominance of an antithesis, it can be classified as either negative, which corresponds to a continuation of the antithesis dominance pending the arrival of a fresh thesis or positive, which corresponds to the arrival of a new thesis. For a given word under the dominance of an antithesis, the probability of it being negative is given by the probability  $q$ , while the probability that it is being positive is given by the probability  $p$ .

Thus, given a thesis occurs, then the thesis length has a probability distribution

$$\Pr[X_i = n] = qp^{n-1} \quad n = 1, 2, \dots$$

This can be seen as follows. Given the occurrence of a thesis, then it must consist of at least one positive word (the first word of the thesis), for otherwise, it would not be a thesis. For long passages, the number of words is large and may be mathematically approximated by infinity. Following the same reasoning, we also have

$$\Pr[Y_i = n] = pq^{n-1} \quad n = 1, 2, \dots$$

The probability generating function  $F(z)$  of  $X_i$  is therefore given by

$$F(z) = \sum_{k=0}^{\infty} \Pr[X_i = k] z^k = \frac{q}{p} \sum_{k=0}^{\infty} p^k z^k \\ = \frac{q}{p(1-pz)}.$$

The probability generating function  $G(z)$  of  $Y_i$  is given by

$$G(z) = \sum_{k=0}^{\infty} \Pr[Y_i = k] z^k = \frac{p}{q} \sum_{k=0}^{\infty} q^k z^k \\ = \frac{p}{q(1-qz)}.$$

The mean thesis length is obtained by differentiation

$$E(X_i) = F'(1) = \frac{1}{q},$$

and similarly, the mean antithesis length is

$$E(Y_i) = G'(1) = \frac{1}{p}.$$

The variance of the thesis length is

$$\text{Var}(X_i) = F''(1) + F'(1) - F'(1)^2 = \frac{p}{q^2},$$

and the variance of the antithesis length is

$$\text{Var}(Y_i) = G''(1) + G'(1) - G'(1)^2 = \frac{q}{p^2}.$$

Another common way of characterizing a passage is determining the total length  $L$  of words in relation to a given length  $r$  of the thesis. Now,  $L=k$  ( $\geq r$ ), iff the  $k$ th word encountered coincides with the  $r$ th positive word, and this happens with probability

$$\Pr[L = k] = \binom{k-1}{r-1} p^r q^{k-r}$$

for  $k = r, r+1, r+2, \dots$ . This can be seen by noting that in order for  $L=k$  to be true, we must have  $(r-1)$  positive words among the first  $(k-1)$  words encountered, and this has the binomial distribution

$$\binom{k-1}{r-1} p^{r-1} q^{k-r}$$

On multiplying this by the probability that the  $k$ th word is positive, we obtain  $\Pr[L=k]$ . The mean value of  $L$  is given by

$$E[L] = r \left(1 + \frac{q}{p}\right).$$

On identifying  $1/q$  as the mean thesis length and that, similarly,  $1/p$  as the mean antithesis length, we obtain the following relation

$$E[L] = r \left(1 + \frac{\text{mean length of antithesis}}{\text{mean length of thesis}}\right).$$

Or in normalizing by  $r$ , we have

$$\frac{E[L]}{r} = 1 + \frac{\text{mean length of antithesis}}{\text{mean length of thesis}},$$

which can be utilized to characterize the style of a passage.

Another useful characterization is represented by the renewal function [8], which in the present context is the number of distinct thesis episodes  $N$  within  $r$  positive words. The  $r$  positive words have at most  $(r-1)$  gaps among them, since between any two successive positive words, there may or may not be intervening negative words. The probability of having no such intervening words is of course  $p$  and that of having one or more such words is  $q$ . Hence, the number of distinct episodes of the thesis equals  $k$  iff there are  $(k-1)$  interventions among the  $(r-1)$  gaps, i.e.,

$$\Pr[N = k] = \binom{r-1}{k-1} q^{k-1} p^{r-k}$$

where  $1 \leq k \leq r$ . This has mean

$$E[N] = 1 + (r-1)q.$$

If  $r$  is large and  $q$  small, then we can use the Poisson approximation

$$\Pr[N = k] \cong \frac{[(r-1)q]^{k-1} e^{-(r-1)q}}{(k-1)!}.$$

## V. EXPERIMENTS AND ILLUSTRATIONS

Leo Tolstoy's novel "War and Peace" and Barack Obama's speech "Yes We Can" are two works that belong to vastly different genres. "War and Peace" is a classic novel that was first published in 1869 and is considered one of the greatest literary works in history. It is a work of historical fiction set against the backdrop of the Napoleonic Wars and explores themes of power, war, love, and society in 19th-century Russia. On the other hand, Barack Obama's speech "Yes We Can" is a political speech delivered by the former US President during his 2008 presidential campaign. It is a work of oratory and aims to persuade and motivate the American people to vote for Obama and support his vision for the country's future. These two works may be vastly different in their genres, but both demonstrate the power of language and storytelling to explore important themes and ideas that resonate with readers and audiences.

### A. Russian Novel

The first chapter of Leo Tolstoy's "War and Peace" sets the stage for the novel's exploration of power, war, and social norms in 19th-century Russia. Chapter One begins with a discussion of the state of Russian society in the early 1800s, with a focus on the importance of social hierarchy and wealth. The reader is introduced to several aristocratic families and their relationships with each other, including the Bolkonskys, the Rostovs, and the Bezukhovs. When Anna Pavlovna talked about family members, Prince Vasili said,

"What would you have me do?" he said at last. "You know I did all a father could for their education, and they have both turned out fools. Hippolyte is at least a **quiet** fool, but Anatole is an **active** one. That is the only difference between them." He said this smiling in a way more natural and **animated** than usual so that the wrinkles round his mouth very clearly revealed something unexpectedly coarse and **unpleasant**.

Therefore, we have

X      2  
Hippolyte is at least a **quiet** fool,  
3   4   5   6   -X  
but Anatole is an **active** one.

He said this smiling in a way  
X      2      3  
more natural and **animated** than usual,  
4   5   6   7   8   9   10  
so that the wrinkles round his mouth  
11   12   13   14  
very clearly revealed something  
15                      16   17   -X  
unexpectedly coarse and **unpleasant**.

The first “X” is “quiet” and its follow-up “-X” is “active”. The length between “quiet” and “active” is 6 words. Then we have the second pair of “X” (animated) and “-X” (unpleasant), and the length is 17 words. The average length of these two sentences is  $(6+17)/2=11.5$  words. Compared with Shakespeare’s “Henry V,” Leo Tolstoy’s “War and Peace” has a shorter length of antithesis. To summarize, we have:

$$\begin{aligned} E(X_i) &= 11.5 \\ E(Y_i) &= 12 \\ \text{For } r = 10, E(L) &= 21 \end{aligned}$$

### B. Presidential Speech

Barack Obama delivered his New Hampshire Primary Concession Speech entitled “Yes We Can” on January 8th, 2008, in Nashua, New Hampshire. The speech was delivered in the aftermath of his defeat in the New Hampshire Democratic primary by Hillary Clinton. Obama had previously won the Iowa caucus, but his defeat in New Hampshire was seen as a major setback for his presidential campaign. The primary was a crucial moment in the race for the Democratic nomination, and Obama’s defeat was unexpected, given the momentum he had gained after Iowa. Obama’s concession speech was a pivotal moment in his campaign, one that would ultimately lead him to win the presidency. In the speech, Obama addressed his supporters, acknowledging the difficult road ahead while rallying them to continue the fight of his campaign. The speech was widely praised for its emotional appeal and demonstration of Obama’s resilience and determination. It is often seen as a turning point in his campaign, ultimately leading to his victory in the Democratic primaries and his ultimate election as the 44th President of the United States. In the speech, Obama goes,

“And whether we are **rich** or **poor**, **black** or **white**, Latino or Asian, whether we hail from Iowa or New Hampshire, Nevada or South Carolina, we are ready to take this country in a fundamentally new direction.

.....

We can bring **doctors** and **patients**, **workers** and **businesses**, **Democrats** and **Republicans** together, and we can tell the drug and insurance industry that, while they get a seat at the table, they don’t get to buy every chair, not this time, not now.”

Therefore, we have the following antithesis:

X   2   -X   X   2   -X

And whether we are **rich** or **poor**, **black** or **white**

X   2   -X

We can bring **doctors** and **patients**,

X   2   -X

**workers** and **businesses**,

X   2   -X

**Democrats** and **Republicans** together

$$\begin{aligned} E(X_i) &= 2 \\ E(Y_i) &= 1+4+1+1=1.75 \\ \text{For } r = 10, E(L) &= 17 \end{aligned}$$

### C. Shakespearean Classic

X

What’s he that wishes so?

-X                  X                  -X

My cousin, Westmoreland? No, my fair cousin;

$$\begin{aligned} E(X_i) &= (5+1)/2 = 3 \\ E(Y_i) &= 2 \\ \text{For } r = 6, E(L) &= 8 \end{aligned}$$

We see that  $E(X_i)$ ,  $E(Y_i) = 2$ , and  $E(L)$  for these three categories of literary styles are distinctly different.

The antithesis is a literary device in which opposites are put close to one another in a sentence or phrase for contrasting effects. In the opening of Shakespeare’s “Henry V” (before Act 1 Scene I), there are several examples of antithesis when Chorus enters and delivers the following speech.

Stanza 1:

O for a Muse of fire, that would ascend

X   2   3

The brightest **heaven** of invention,

4   -X                                  X

A **kingdom** for a stage, princes to **act**

2   3                  4   -X

And monarchs to **behold** the swelling scene!

Then should the warlike Harry, like himself,

Assume the port of Mars; and at his heels,

Leash’d in like hounds, should famine, sword and fire

Crouch for employment.

But pardon, and gentles all,

X

The flat unraised spirits that have **dared**

2   3   4                  5                  6   7   8

On this unworthy scaffold to bring forth

9   10   11   12                  13   14   15   16

So great an object: can this cockpit hold

17   18   19   20   21                  22   23   24   25

The vasty fields of France? or may we cram

26   27   28                  29   30   31   32

Within this wooden O the very casques

33   34   -X

That did **affright** the air at Agincourt?

There are at least four pairs of thesis and antithesis in the above literary passages, such as heaven and kingdom, act and behold, scaffold and cockpit, dared and affright. These examples of antithesis in stanza 1 serve to emphasize the

tension in Shakespeare's play, particularly in regard to the relationship between the grand ambitions of artistic creation and the limitations of the physical world.

First, the length between "heaven" and "kingdom" is 4. Regarding heaven and kingdom, the thesis in this contrast is the idea that the kingdom is like a stage, a place of grandeur and excitement worthy of the presence of princes and monarchs. The stage is an unworthy scaffold, a place that is limited in scope and unable to contain the vastness of the world outside like heaven. However, the antithesis of the kingdom is heaven, which is limitless because it usually refers to the expanse in which the celestial bodies, such as the sun, moon, and stars are observed, which was historically considered to resemble a large vault stretching over the earth, also known as the sky or the firmament.

Second, in terms of "act" and "behold", the length between them is 4 as well. The thesis is that the stage is a place for actors to take on roles, an opportunity to "act" and perform for audiences. The antithesis is that the monarchs and audience members who watch the play are passive spectators who "behold" the action, emphasizing their potential power and leashed engagement.

Third, the length between "dared" and "affright" is 34 words, which is largely distinct from the previous two pairs of thesis and antithesis. According to [11], "affright" means "to frighten, terrify" and "to be or become afraid", while "dare" refers to "having boldness or courage". The thesis here is that the actors have dared to bring forth an object of great significance, symbolized by the "great object" they are attempting to create through their performance. The antithesis is that the objects, such as the "casques" (or helmets) used in the battle scene which "affright" the air at Agincourt, are so terrifying that they cannot be contained within the "wooden O" symbolizing the limited confines.

Stanza 2:

O, pardon! since a crooked figure may  
                   X   2   3   4  
 Attest in **little** place a million;  
 5   6   7   8   9   10 -X  
 And let us, ciphers to this **great** accompt,  
 On your imaginary forces work.  
 Suppose within the girdle of these walls  
 Are now confined two mighty monarchies,  
 Whose high upreared and abutting fronts  
                   X   2   3  
 The perilous narrow **ocean** parts asunder:  
 4   5   6   7                   8   9   10  
 Piece out our imperfections with your thoughts;  
 11 12 13   14 15 16 17  
 Into a thousand parts divide on man,  
 18 19 20           21  
 And make imaginary puissance;  
 22 23 24 25 26 27   28 29 30 31  
 Think when we talk of horses, that you see them

32       33   34       35 36 37   38   -X  
 Printing their proud hoofs i' the receiving **earth**;  
 For 'tis your thoughts that now must deck our kings,  
                   X   2   -X

Carry them **here** and **there**; jumping o'er times,  
 Turning the accomplishment of many years  
 Into an hour-glass: for the which supply,  
 Admit me Chorus to this history;  
 Who prologue-like your humble patience pray,  
 Gently to hear, kindly to judge, our play.

The thesis and antithesis in the passage are ocean and earth, little and great, here and there. These antithesis and thesis emphasize the grand ambition of the playwright, who attempts to transport the audience's imagination from the small confines of the stage to vast distances and times. The use of antithesis serves as a reminder of the limitations that can constrain even the most ambitious attempts to capture the epic grandeur of history on the stage.

Between "little" and "great", the length is 10 words. Here, the thesis is that a small place can be of great importance, while the antithesis is that something great can also be imperfect and have flaws. The line "a crooked figure may / Attest in little place a million" highlights the significance of small things in the grand scheme of things.

Regarding "ocean" and "earth", the length is 38 words. The thesis is the idea that two mighty monarchies are separated by the narrow and perilous "ocean" that lies between them. It suggests a vastness and distance that capture the imagination, allowing the audience to visualize and join the drama. The antithesis of "earth" is that these monarchies are confined within the walls of the stage, reducing the scale of the conflict and the drama.

"Here" and "there" share a relatively shortest length of 2 words. The thesis aims to transport the audience's imagination across vast distances and times, from "here" to "there", referring to the distant monarchies and their battles. The line "Carry them here and there; jumping o'er times" emphasizes the idea of travelling through time and space. The antithesis is more subtle but can be seen in the line "Piece out our imperfections with your thoughts," which suggests that imagination is needed to fill in the gaps and make the performance more complete.

These antitheses in the above two stanzas highlight the grandeur and magnitude of the events depicted in the play and contrast the lofty aspirations of the characters with the limitations of the stage and the mortal world. The use of antithesis also adds rhetorical complexity and depth to the language, highlighting Shakespeare's skill as a writer and his ability to convey powerful emotions through his use of language.

## VI. SUMMARY AND CONCLUSION

Antithetical elements are often used by authors in English literary writings and speeches. Antithesis refers to the juxtaposition of contrasting words or ideas, which often, although not always, in parallel structure. Such rhetorical

figure is used to contrast opposing ideas, creating a sense of tension and urgency, as well as heightening the emotional impact of the speech. Contrasting the antithesis is the thesis, which refers to a statement, assertion, or tenet. Thesis also is a proposition laid down or stated, especially as a theme to be discussed and proved, or to be maintained against attack. It is observed that the points of occurrence of thesis and antithesis are random, yet they occur with remarkable regularity alternating each other. In this study, we represent the lengths of the thesis and antithesis by using an alternating renewal stochastic process. We find that the underlying attributes, such as the length of the thesis, the length of the antithesis, and the renewal function, of the alternating renewal process are able to usefully characterize the writing style of individual authors and help us to quantify and understand their linguistic technique and intention. Experiments are carried out to illustrate the present approach. Using the present method in conjunction with other techniques, such as sentiment analysis, and word embedding models, it is possible to gain a deeper understanding of the literature and its underlying themes and structures.

#### REFERENCES

- [1] C. H. C. Leung and C. Zeng, "The Use of Multi-Step Markov Chains in the Characterization of English Literary Works," in *Proceedings of the 11th International Conference on Data Analytics*, pp. 43-48, 2022.
- [2] C. Zeng and C. Leung, "The Use of Stochastic Models in the Analysis of Vast English Literary Data Corpora," in *2020 6th International Conference on Big Data and Information Analytics (BigDIA)*, pp. 282-288, 2020.
- [3] C. H. C. Leung and C. Zeng, "Pattern Discovery and Stylometric Analysis in English Literature and Literary Translation Through State Integration in Markovian Representations," in *Journal on Advances in Software*, vol. 16, no. 1&2, June 2023, in press.
- [4] E. Smith. *The Cambridge Introduction to Shakespeare*. Cambridge University Press, 2007.
- [5] W. Shakespeare. *Henry V*. Oxford University Press, 2008.
- [6] M. Eder, M. Kestemont, and J. Rybicki, "Stylometry with R: A package for computational text analysis," in *R Journal*, 8(1), pp. 107-121, 2016.
- [7] D. K. Jha, "Markov Modeling of Time-Series Data using Symbolic Analysis," in *arXiv Labs*, 2021. Available from: <https://doi.org/10.48550/arXiv.2103.11238>
- [8] R. Bass. *Stochastic Processes*. Cambridge University Press, 2012.
- [9] L. Tolstoy. *War and Peace*. Duke Classics, 2013.
- [10] B. Obama. Yes We Can, 2008 Presidential Campaign Speech.
- [11] Oxford English Dictionary.
- [12] P. Honan. *Shakespeare: A Life*. Oxford University Press, 1999.
- [13] C. L. Barber. *Shakespeare's Festive Comedy*. Princeton University Press, 2011.
- [14] R. N. Raximovna and M. M. Mirusmanovna, "Comparison of Pragmatic Value of Antithesis in Fiction Texts," *Turkish Journal of Computer and Mathematics Education* 12.6, pp. 1195-198, 2021.
- [15] V. S. Bailey, "Pope and Antithesis: "Law and War with Words,"" in *Studies in English Literature 1500-1900*, 27(3), pp. 437-454, 1987.
- [16] M. Dobson, W. Wells, W. Sharpe, and E. Sullivan. *The Oxford Companion to Shakespeare*. Oxford University Press, 2016.