

## Linguistic Evidence of the Theme of Nature in a Constellation of Robert Frost's Poems

Amira Hasan Ahmed <sup>a</sup>

[ameera.saad@yahoo.com](mailto:ameera.saad@yahoo.com)

Universiti Malaysia Sarawak, Malaysia

Monaliza Sabrini-Zin <sup>b</sup>

[szmonaliza@unimas.my](mailto:szmonaliza@unimas.my)

Universiti Malaysia Sarawak, Malaysia

### ABSTRACT

While Robert Frost is quite known for being a poet who wrote about nature, especially rural areas of New England in America, on trees, dark forests and nights, apples and apple harvesting, skies, stars, and even roads not taken, the use of such elements of nature in his poems and its contribution to the theme has not yet been adequately and collectively addressed. The current study focuses on analyzing a group of poems written by Robert Frost using AntConc and Lancbox (corpus linguistic toolkits) to provide linguistic evidence of the existence of the nature theme. This study examined how keywords, collocation, and colligation from different poems displayed the theme of nature. Based on a corpus stylistic approach, the study found that analyzing poems from different volumes shows the impact of the environment on the poet was positively and negatively associated with the poet's convictions with nature. Robert Frost frequently clashes with nature in his themes of acceptance, loneliness, and terror. The poet's exposure to the rural landscape led him to appropriate many pictures from his surroundings and relate them to religious and universal ideas that form the core of his poems. In addition, because of its relatability and simplicity, Robert Frost's poetry has profoundly impacted many readers. This paper contributes to the current literature by revealing that Robert Frost's poems portray nature, showing a connection between humanity and nature and emphasizing the impact of nature on people's attitudes.

**Keywords:** Robert Frost; corpus linguistics; reference corpus; keywords; collocations

### INTRODUCTION

Robert Lee Frost (1874-1963), (henceforth, RF) enjoyed widespread popularity in America; he was the most taught and read in American literature and the globe. He received the Pulitzer Prize four times and was the only poet ever invited to read his poem, "The Gift Outright," at the inauguration ceremony of President Kennedy in 1961; he was the only poet who received such an honor. Undoubtedly, all these honors made him the poet laureate in American literature. RF's modern poetry, influenced by youth, features traditional verse form, clear sentences, and language, showcasing a connection between man and nature, making it a valuable resource for readers.

---

<sup>a</sup> Main author

<sup>b</sup> Corresponding author

Robert Frost's life and literary legacy reflect a journey of personal resilience and creative brilliance. Born in San Francisco and named after Confederate General Robert E. Lee, Frost experienced significant change early in life, moving to eastern Massachusetts with his mother and sister after his father's death. As the son of a prudent and hard-working Massachusetts farmer, Frost exhibited remarkable creativity and productivity, leaving behind an impressive body of poetry. Among his celebrated works are *A Boy's Will*, *Mountain Interval*, and *New Hampshire*, which continue to captivate readers worldwide.

Contemporary approaches are increasingly used to study linguistic features in nature-themed poetry, particularly RF's works, as traditional stylistic methods become outdated and poetry remains challenging to analyze. According to Verdonk (2013), poetic texts are inherently intricate because readers often produce diverse interpretations of the same text. Using a corpus stylistic approach, this study examines a selection of RF's poems from PG to explore his subtle and meaningful depiction of nature. RF's keen observations of New England's rural landscapes are evident in his detailed portrayal of small natural elements as representations of the whole. His quiet and modest expression of a deep connection to nature, described by Adams (1965) as both embracing and accepting, highlights the richness of his poetic style. This research contributes to corpus linguistics (CL) while offering fresh insights into RF's engagement with nature.

RF's On the other hand, readers can be aware of the existence of the natural elements in the poems, but this study can provide better critical reading ability. This study presents evidence about how our intuitive knowledge about well-known poems can be supported by statistics utilizing corpus linguistic toolkits. Most importantly, the study can help realize these representations through linguistic evidence and contribute to a more inclusive understanding of pastoral words/ natural elements in RF's poems. This study aims to answer three questions namely, 1) How is nature represented in RF's poems? 2) Have keywords, collocations, and colligations contributed to the theme of nature? and 3) Is Corpus stylistic approach a successful method to reveal the theme of nature?

## PAST STUDIES

As RF agrees with the view of human existence in the natural world, Montgomery (1958, p. 339) noted that "his attitude toward nature is one of the armed and amicable truces and mutual respect interspersed with crossings of the boundaries separating the two principles, individual man and force of the world. But boundaries are insisted open" (After reading his poems, Untermeyer (1920) admits the reality and power of his work. RF is not the dark naturalist that many believe he is. Behind the mask of "grimness" that many of his critics have placed on him, there is a constant elfin pucker; a whimsical smile, half-disclosed raillery, glints beneath his most serious monologues. RF's most concrete facts are symbols of spiritual values (Untermeyer, 1920).

Despite his concern with the natural world, he honestly portrays the complexity of human experience. The symbols of his poems like 'Birches' and 'Wild Flower', 'Wood' and 'Stones Walls', 'Pasture Spring' and 'The Snow of Northern England' are recognized locales from the days of his writing career. If the rural images and symbols from his works are removed, we cannot identify him as a vigorous writer, or absolute genius: successful in speech and meaning in the world of nature. The most widely held misconception about him may be the tendency to regard him as a nature poet exclusively.

RF's poems are realistic and transparent, focusing on rural life and nature. They convey messages about life stages but differ from Wordsworth's emotional depiction of the natural world. RF's poetry reflects a love and truce between man and nature, crossing boundaries. On the basic difference between Wordsworth and RF, Sharma (2011) states:

Certainly, RF is not a poet of Nature like Wordsworth is. To some extent, Nature is RF's subject. But in his poetry, Nature is only a background; it is a backdrop behind the humming scenes of human beings at work. RF has been excessively hailed as a poet of Nature even when he once confessed that he was not a Nature poet. It is said that RF was influenced by Wordsworth but there is very little information about the nature of the influence. (Sharma, 2011, p. 55-59)

For RF, the task of poetry is the reactivation of the connection between word and world, the transformation of the “grammatical sentence” into the “vital sentence” (Hass, 2005). Moreover, it requires that we credit, however, provisionally, our intuitive understanding. In Almiqdady et al. (2015), for instance, the concept of nature is discussed in the poetry of RF and Wordsworth comparatively. The research is important because it sheds light on the similarities and differences between the two poets in terms of their attitudes towards nature to familiarize the readers with the poetic styles of both poets and their view regarding nature as a historically known source of motivation and aspiration.

Additionally, numerous works and studies have been published on the relationship between RF and Nature such as Rashid et al. (2012). Critics (such as Baym, 1965; Lynen, 1960; Naess, 2015; and Zeng, 2024) have labeled RF as a ‘nature poet’ and then assumed that he was a version of Emerson or Wordsworth as though there was only one way to be a Nature poet. Nature in RF's view represents spirit. As he publishes something about nature, he does not explain it but instead utilizes real phenomena or occurrences as metaphors to expose and convey something more meaningful. (Nitchie, George, 1978) reflects on RF's understanding of reality, noting that RF's ideals and experience of reality are closely linked. This study identifies RF's poems' connection to nature, natural elements, and setting, but differs in methodology for assessing nature's contribution to themes and shaping RF's unity.

## CORPUS-BASED STYLISTIC APPROACH

Despite for a long time enjoying only a peripheral status within stylistics and literary studies, the practice of corpus stylistics has grown in recent years, to the extent that it is fast becoming a recognizable field within stylistics generally (McIntyre & Walker, 2019). The corpus-based approach is the use of statistics of large collections of written or spoken data (corpora) to investigate linguistic phenomena. Researchers in the field who show good examples of the application of corpus stylistic approaches on literary and non-literary works include Mahlberg, 2013a, McIntyre, 2012, Fischer-Starke, 2010, Toolan, 2009, Ho, 2011, Mahlberg, 2013b, Hoover et al. 2014 and Demjen, 2015.

Through large collections of written or spoken language in machine-readable form, various issues concerning form and meaning can be examined. The applications of corpora involve both quantitative and qualitative methods to gather naturally occurring language using computational tools and software such as a concordance, a software, either installed on a computer or accessed through a website, which can be used to search, access, and analyze language from a corpus; in our case it is installed on the computer. Based on large collections of real-life language examples, three major outcomes are made available out of these tools: *frequency lists*, *concordance lines*,

and *collocations*. According to Firth (1957), collocations are words that go together. Frequency and concordance lines provide empirical bases for the qualitative study of a wide range of linguistic and extra-linguistic phenomena. For instance, studying collocation is accessible through the co-occurrence of particular words with other words in many examples as validated by concordance lines. The recurring patterns in corpora allow researchers to support their qualitative analysis and findings. Corpus linguistics methods are powerful means to uncover hidden patterns of language use that are not observable to researchers without computer aid.

Many studies tackled literary texts from a corpus-based perspective. For example, Murphy (2006) used this approach to analyze soliloquies and asides in 12 Shakespeare plays. His study is based on keywords, grammatical categories, and semantic field analysis in these plays. Similarly, in recent years, there are corpus-based studies that have also examined aspects such as differences between female and male language (Hota & Koppel, 2006), key semantic domains and metaphors in Shakespearean love tragedies and love comedies (Archer et al., 2005), characterization in 'Romeo and Juliet' (Culpepper, 2001), and imagery in Macbeth (Zyngier, 1999).

Like other research approaches, corpus linguistics techniques are not perfect. Researchers have voiced their concerns and objections to the use of corpora. Early objections against the usage of corpora were enumerated by McEnery and Wilson (2001) in three points: 1) Its competence should be modeled by introspection in studying linguistics, not performance as encouraged by studying a corpus. 2) Because natural language is infinite 2). Since natural language is infinite, describing performance is still unattainable and, thus, inadequate and 3) introspection must not be avoided, so that ungrammatical and ambiguous structures are possible. Widdowson (2000) and other academics criticized the discrepancy between the findings and the intuitions of native speakers. Borsley and Ingham (2002) used the challenge of accounting for absent or missing cases in a corpus as another justification for rejecting corpus approaches.

The theoretical framework is a corpus stylistic approach as in McIntyre and Walker (2019), focusing on the lexical features in the poems using the tools in the AntConc (Anthony, 2021), a corpus linguistic software. This approach is naturally quantitative, which is increasingly being used in applied linguistics (McIntyre & Brian, 2019). This paper utilizes a specialized corpus consisting of some selected RF's poems, which was built to study lexical features of language by searching for highly frequent patterns (i.e., recurrent clusters of certain sets of lexical choices). The software AntConc (version 4.2.0) (Anthony, 2022) was used to investigate lexical clusters, collocates, and their concordances to identify the dominant semantic macrostructures of the natural elements involved in shaping concepts concerning the theme of nature in RF's poems. The integrated concordancer allows computer-aided careful inspection of each word form for qualitative analysis. Concordancers also pinpoint the text where the search term occurs (listing the file name), which can be useful for the study of the distribution of a lexical unit or for identifying a particular set of texts for closer analysis. Hence, it is possible to meaningfully combine quantitative and qualitative analyses, using frequency and "keyness" analyses. The results retrieved from the software are followed up by qualitative text analysis. Here, it should be noted that the concordancer significantly aids the manual and qualitative analysis of the data: the software automatically provides all instances for particular search terms and further information on such terms (Anthony, 2012; Bednarek & Carr, 2020). Moreover, another tool is used in this study, Lanksbox. The graphcoll tool in Lanksbox (0.6) identifies collocations and displays them in a table and a collocation graph or network.

Identifying keywords is one of the crucial techniques in corpus linguistics (Scott, 1997 as cited in Brezina, 2018). Keywords are words that are considerably more frequent in one corpus (node corpus, target corpus, or corpus of interest) than in another corpus (reference corpus). In other words, “keywords” is a relative term depending on the differences in lexical frequencies in the two corpora in question. In this study, the target corpus is the selected poems downloaded from Project Gutenberg (henceforth, PG) while the reference corpus as selected here is a sub-category within the Brown Corpus, i.e., Am06 (PRESS REPORTAGE) which has been downloaded from Anthony’s software as it already exists. The comparison is based on a statistical measure which is defined by using different measures of keyness. Keyness can be computed for words occurring in a target corpus by comparing their frequencies (in the target corpus) to those in a reference corpus. The keyness method in this study, AntConc), however, provides us with the most specific vocabulary for the particular type of text by counting not the simple frequency, but rather the unusual frequency as compared to the reference corpus. In practice, this is done automatically by the corpus software. It is important to note that the reliability and relevance of comparisons are enhanced when the reference corpus is both larger and more similar to the corpus of interest. As Baker et al. (2006) highlight, the selection of a comparable reference corpus is critical to ensuring meaningful and accurate analysis, as it provides a robust basis for identifying significant linguistic patterns and variations.

This study aims to generate a corpus of RF's poems from PG, ensuring copyright protection and easy retrieval of digital content, despite potential copyright violations. The selected poems were mainly taken from three volumes: Mountain Intervals (1916), North of Boston (1914), A Boy’s Will (1913), and one poem from New Hampshire (1923) which is The Runaway. No instances outside the data-defined scope of RF’s poems were included, meaning no other poet’s poems were added (Table 1). All the poems under study were saved in (txt. File) format that is compatible with AntConc and easy to use in the corpus software.

TABLE 1. List of RF’s poem texts under study

S/N	Title of the Poems	Volume Taken from	Date of Publication
1	THE PASTURE	North of Boston	1914
2	THE COW IN APPLE-TIME	Mountain Interval	1916
3	THE RUNAWAY	New Hampshire	1923
4	AN OLD MAN'S WINTER NIGHT	Mountain Interval	1916
5	HOME BURIAL	North of Boston	1914
6	THE DEATH OF THE HIRED MAN	North of Boston	1914
7	A SERVANT TO SERVANTS	North of Boston	1914
8	THE SELF-SEEKER	North of Boston	1914
9	THE HILL WIFE	Mountain Interval	1916
10	"OUT, OUT. . ."	Mountain Interval	1916
11	PUTTING IN THE SEED	Mountain Interval	1916
12	GOING FOR WATER	A Boy’s Will	1913

---

13	MOWING	A Boy's Will	1913
14	AFTER APPLE-PICKING	North of Boston	1914
15	BIRCHES	Mountain Interval	1916
16	THE GUM-GATHERER	Mountain Interval	1916
17	THE MOUNTAIN	North of Boston	1914
18	THE TUFT OF FLOWERS	A Boy's Will	1913
19	MENDING WALL	North of Boston	1914
20	AN ENCOUNTER	Mountain Interval	1916
21	THE WOOD-PILE	North of Boston	1914
22	SNOW	Mountain Interval	1916
23	IN THE HOME STRETCH	Mountain Interval	1916
24	THE ROAD NOT TAKEN	Mountain Interval	1916
25	THE OVEN BIRD	Mountain Interval	1916
26	A VANTAGE POINT	A Boy's Will	1913
27	THE SOUND OF TREES	Mountain Interval	1916
28	HYLA BROOK	Mountain Interval	1916
29	RANGE-FINDING	Mountain Interval	1916
30	OCTOBER	A Boy's Will	1913
31	TO THE THAWING WIND	A Boy's Will	1913
32	A TIME TO TALK	Mountain Interval	1916
33	THE CODE	North of Boston	1914
34	A HUNDRED COLLARS	North of Boston	1914
35	BLUEBERRIES	North of Boston	1914
36	BROWN'S DESCENT	Mountain Interval	1916
37	REVELATION	A Boy's Will	1913
38	STORM-FEAR	A Boy's Will	1913
39	BOND AND FREE	Mountain Interval	1916
40	FLOWER-GATHERING	A Boy's Will	1913
41	RELUCTANCE	A Boy's Will	1913
42	INTO MY OWN	A Boy's Will	1913

---

## DATA ANALYSIS

As previously discussed, the data for this study, comprising selected poems by RF, were obtained from PG. PG is an online library of free electronic books or eBooks. PG is a voluntary effort to digitize, archive, and distribute literary works. This online content provider has an extensive history and continues functioning, thus remaining one of the oldest in existence. PG supports and publishes many open formats. There are more than 70,000 documents on the website ([www.gutenberg.org](http://www.gutenberg.org)) which is a digital platform that offers literary works in various formats. The releases are available in plain text (txt. File) as well as other formats, such as HTML, PDF, EPUB, MOBI, and Plucker wherever possible. In this study, the format which is used here is plain text (txt. File) as it goes well with AntConc software. Most releases are in the English language, but many non-English works are also available. PG only releases works that are in the public domain in the United States, and it claims no copyright at all in its releases. The data for this study consists of over **45488** tokens (words). It is not the researcher's choice to select the following poems for a neutral discussion. The notable feature that distinguishes the content obtained from the Gutenberg website is its exemption from copyright restrictions. The selected poems fulfill the study's overarching aim of investigating linguistic evidence through lexical choices and features. We examined the poems to showcase how Nature has been tackled in each poem and what the differences are.

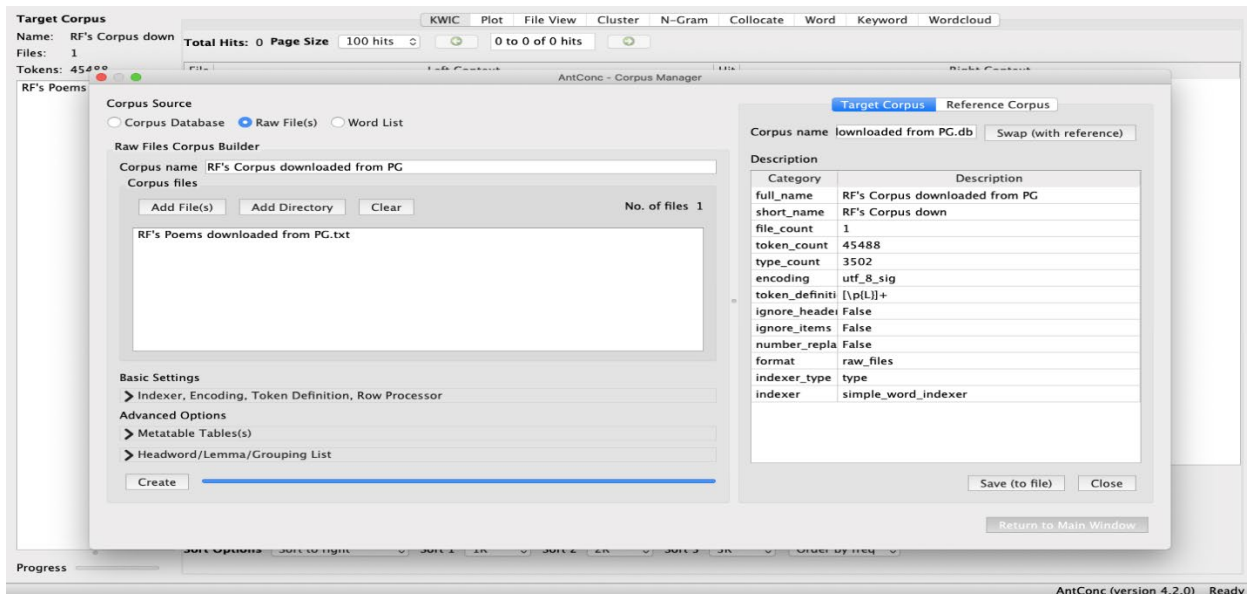


FIGURE 1. Uploading the node/ target corpus (RF's collected poems)

The analyses of the data collected from the PG (node) are used to obtain quantitative and qualitative concordance lines and to provide a comprehensive analysis of the natural elements in RF's poems and how these elements contribute to the theme.

Therefore, the methodology used here is both qualitative analysis and quantitative analyses are combined. Quantitatively, corpus-based analysis uncovers the facts within the corpus language patterns which in turn can be interpreted through keywords, collocation, concordance lines, and co-texts. Qualitatively, reading the statistics collected from AntConc and interpreting them is complementary to the quantitative method. Thus, employing corpus-based analysis on the poems

can help reduce analysts' bias, help them obtain a much better picture surrounding the frequency of particular phenomena (Baker et al., 2013), and enable them to go beyond single texts and gain insight into the Nature theme in RF's poems.

Keywords are important when identifying key concepts in discourses, typical vocabulary in a genre/language variety, lexical development over time, etc. (Brezina, 2018). Complementary to keywords are lockwords, a term introduced by Baker (2011). Lockwords are words that occur with similar frequencies in two corpora that we compare. The importance of keywords (lockwords) lies on our ability to compare two corpora in a meaningful way.

TABLE 2. Decision about keywords: BASIC options

Corpus of interest C	Reference Corpus R	Decision
Frequent	Infrequent	+ (Positive keyword)
Infrequent	Frequent	- (negative keyword)
Comparable freq.	Comparable freq.	0 (lockword)

Moreover, if a word is used more often in C than in R, it is called a positive keyword (+); if, on the other hand, a word is used less in C than in R, it is called negative keyword (-).

Unpacking the Keyword Procedure:

- a) A corpus of interest is RF's poems (with 41 poems in total) as shown in the above screenshot and represented in one file with total number of tokens (45488) and token types (TT= 3502).
- b) Reference corpus is a sub-corpus AmE06 (K=Fiction \_ general. db). The reference corpus is larger in size (59568) with 29 files and with similar in Category. Such heterogeneity between the two corpora can highlight different degrees of in the keyword procedure.
- c) The statistical measure: The log-likelihood statistic that is larger than 3.84 which is the cut-off point (Threshold) for significance at  $p < 0.05$  level (found in the statistical tables or the direct output of the computer program, *AntConc*, as in our study). For both keyword lists, the keyword statistics used was log-likelihood and the threshold was set to  $p < 0.05$ . Log-likelihood is a statistical significance test that verifies the dependability of an observed frequency difference, i.e., how likely it is that a frequency difference is not random. It uses word frequency and corpus size values in its calculations. By setting  $p < 0.05$  (default setting), *AntConc* returns only words that have a log-likelihood score higher than 3.84, which means that there is a 95% chance that the difference is significant. In social sciences, a p-value of 0.05 is generally considered the minimum necessary for confidence in significant results (McIntyre & Walker, 2019). Any values above  $p < 0.05$  means that we are 95% confident that we have a significant result and not one that is simply down to chance (ibid). The effect size measure used was Hardie's Log Ratio, and the threshold was set to all values. According to this measure, if the effect is 0, it means the relative frequency of a word is the same between the corpora being compared. If the value is 1, the word at hand is two times more common in the corpus under study than in the reference corpus. As our study corpus is small, we considered all the words whose effect is above zero in such corpus when compared to the reference corpora.



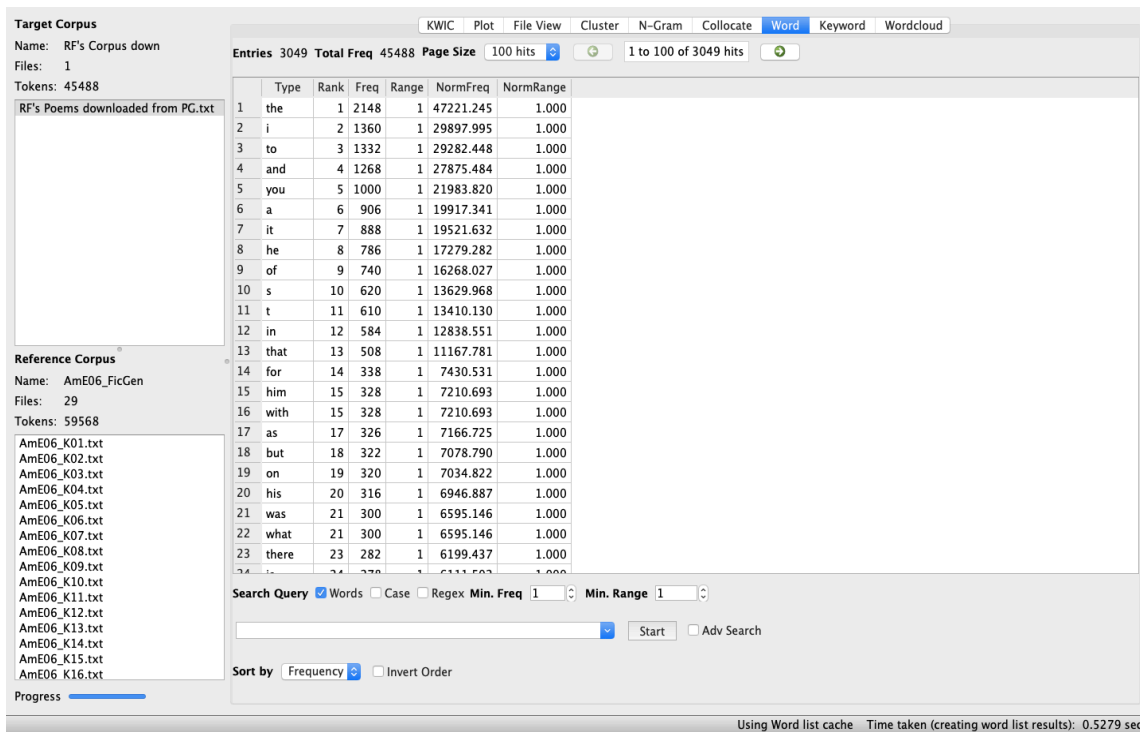


FIGURE 2. Wordlist from RF's selected poems corpus

The list of keywords generated is often very long, and in such cases, it is often an impossible task to analyze each of the keywords in turn. In such cases, we need a means of reducing the amount of data to analyze. One means of doing this is to use a different cut-off point for statistical significance. The corpus linguistic software calculates keyness using a statistical method such as a log-likelihood or chi-square test. Adjusting the p-value in such tests (that is, increasing the level of confidence we want to have concerning whether our findings are indeed statistically significant) will reduce the list of items. Sometimes, this alone is not enough, and we are still left with too much data to handle. In this setting, the comparison of *these poems* to *AmE06* which was stored in AntConc as a reference corpus because it was created using the same sampling frame (Brown family sampling frame) returned a total of 100 keywords (hits).

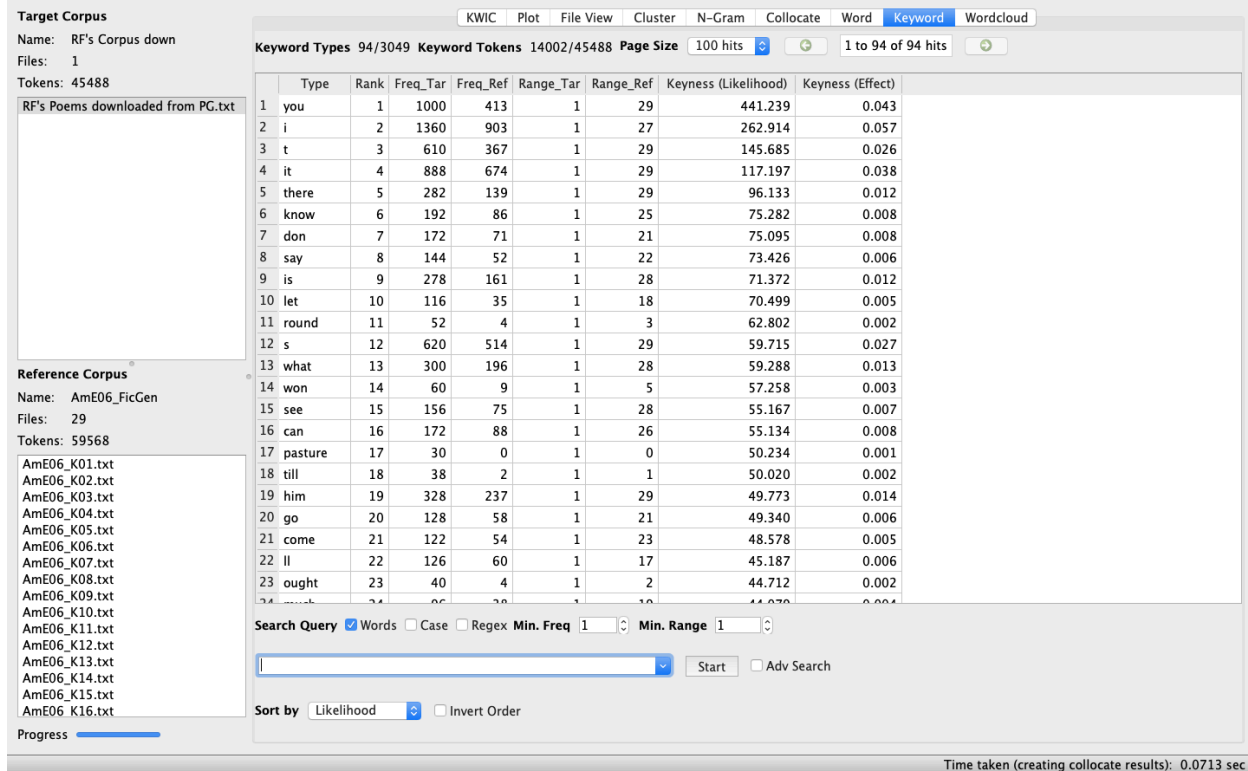


FIGURE 3. Keywords produced by comparing RF’s corpus with Reference Corpus(F-Gn)

For the identification of keywords, Kilgarriff’s (2009) simple maths parameter (hereafter, SMP) was used with the constant 100. Kilgarriff (2009) suggests looking at the ratio between the relative frequencies of words in C and R. Because the ratio can be calculated only if the values in R are greater than zero. Kilgarriff suggests adding constant *k* to both relative frequencies before calculating the ratio. The constant can be any positive number of our choice, but typically it is 1, 10, 100, or 1,000. The simple maths parameter (SMP) is calculated as follows:

$$\text{simple maths parameter (word)} = \frac{\text{relative frequency of a word in C} + k}{\text{relative frequency of a word in R} + k}$$

in this study, the focus was on positive keywords as mentioned previously (more frequent in the corpus of interest **C** than in the reference corpus **R**) further analyzed. These keywords were further verified and analyzed. The constant *k*, we will use here is 100 as we have got 94 hits (keywords):

$$\text{simple maths parameter (you)} = \frac{\text{relative frequency of (you) in C} + k}{\text{relative frequency of (you) in R} + k}$$

$$\text{For example, SMP (you)} = \frac{1000+100}{413+100} = 2.14$$

TABLE 3. Verification of the positive keywords in three ways: log-likelihood, log-ratio, and SMP

Keywords (+)	Keyness (Log-likelihood)	Keyness (Log-Ratio) effect	SMP
1. we	36.836	0.010	1.23
2. you	441.239	0.043	2.14
3. I	262.914	0.057	1.45
4. there	96.133	0.012	1.59
5. know	75.282	0.008	1.56
6. say	73.426	0.006	1.60
7. let	70.499	0.005	1.6
8. go	49.340	0.006	1.44
9. come	48.578	0.005	1.44
10. think	42.533	0.012	1.28
11. mountain	42.533	0.001	1.28
12. night	42.074	0.004	1.40
13. barn	39.319	0.001	1.26
14. wonder	39.319	0.001	1.26
15. make	38.545	0.005	1.37
16. wall	27.717	0.002	1.28
17. glad	24.484	0.001	1.21
18. stove	23.437	0.001	1.16
19. wood	21.814	0.001	1.2
20. way	21.011	0.005	1.20
21. snow	20.155	0.002	1.22
22. something	18.390	0.003	1.21

SMP can be calculated only if the values in R are greater than zero (division by zero is not defined in Mathematics). In Table 3, six keywords, ‘brook’, ‘farm’, ‘flower’, ‘lantern’, ‘lamp’, ‘Cole’, ‘pasture’ and ‘load’ have zero frequency in the reference corpus. Log-likelihood and log-ratio were sufficient parameters in these words.

The extended context of each concordance line is explored to determine its thematic focus on nature in RF’s poems. These keywords are thematically clustered, considering that some keywords feature more than one thematic focus. Eight thematic signals were explored to investigate the theme of nature in the studied poems are organized and projected in the collected data, namely mountain, barn, wall, wood, way, and snow. Figure 3 shows the frequency of concordance lines incorporating the following keywords as displayed below.

From the literature, in several poems, RF adeptly employs nature and its constituent components as a means to articulate his personal experiences, which he has constantly observed. For the top 100 keywords, the keyness score starts very high (at 1360) for the word ‘you’, and gradually decreases to (70) for the word ‘something’. Generally, the keyness of the top 100 keywords vary in a large range. The top five keywords have extremely high keyness, indicating that these five words are necessities for nature-related texts. The keyness decreases from 1000 to 20 with noticeable salience. The keywords with keyness beneath 1000 are common topics that are usually mentioned when some related events take place.

To analyze and refine the selection of keywords, a classification process for the data was pursued. The keywords’ categorization is a fairly subjective process, keywords related to nature were selected. Words such as pronouns, verbs, and modals were excluded. Therefore, few words were left about nature which makes analyzing these words much easier. The data were reduced to

five words only, namely: mountain, barn, wall, wood, way, and snow. Each word is considered a dominant theme by itself with references to other sub-themes.

Table 3 provides a representative overview of the nature theme through keywords from different perspectives to discuss later. First, the collocations in AntConc showed what words accompanied each keyword. Second, using the KWIC (keyword in context) helped in explicating the theme behind every word as shown on the software interface. In addition, scrutinizing the words in collocation helps see the semantic environment around the words and their contribution to the theme.

Nature collocates are nouns, adjectives, or verbs before or after the selected words in RF 's poem corpus, with a one-word window span. Table 4 presents the most frequent collocates of nature-related words. The purpose of this analysis is to discern the link between the node words and their collocates, as well as to determine their collective contribution to the overall theme. Hence, Graphcoll facilitates the visual representation of collocations and their interrelationships (Figure 4).

TABLE 4. Collocates of nature-related words in the corpus

Keywords	Collocate
Mountain	The, strewn, notch, ranges, mass, spruce, land, held, stood, top, pasture, five
Barn	The, comfortable, grows, floor, someone, under, door
Wall	The, within, between, a , builders, mending, outside, run, stone
Wood	Pathless, sweet, the, mid, pile, yellow, bird
Way	the, down, other, long, my, humble, fashioned, lovely, slumped, awkward, nice, ben, station, leads, half, to, its, either, hole.
Snow	The, white, gulped, crusted, looped, settled, crust, storm, bank, hard, help, upon, against

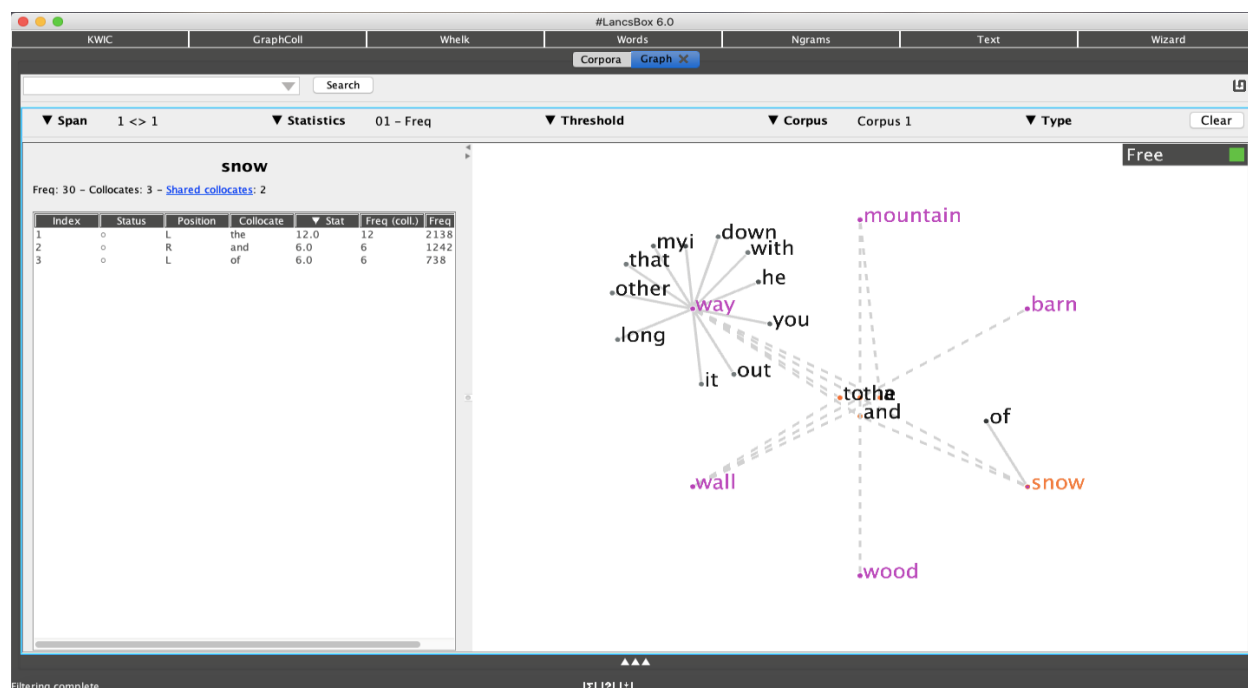


FIGURE 4. Collocations using LancsBox's GraphColl

## RESULTS/ DISCUSSION

By carefully looking at the keyword, it becomes clear that most well-known poems by RF all deal with obstacles. In the selected poems, the barriers are mountains, timber, and snow rather than walls, which are man-made artificial barriers. Obstacles of various kinds, obstacles that separate men from one another, and barriers that appear to separate nature from humans are RF's main concerns as they come through in his poems. In these harsh deserts, man's sport and sports spirit are by these barriers between man and the environment. Nature presents man with these obstacles and difficulties. Man must now demonstrate his mettle and claim to be superior to nature, using his will and mind to control inanimate objects.

**Mountains:** In the corpus of RF, mountains were mentioned 15 times, symbolizing various aspects of human experience. They symbolize awe, strength, endurance, natural beauty, obstacles, perseverance, spiritual growth, enlightenment, and isolation. RF's mountain imagery explores the human experience and highlights the significance of mountains in his broader themes.

RF's depiction of mountains serves as a poignant reminder that pursuing meaningful accomplishments necessitates a willingness to confront personal limitations and surmount adversities. They serve as a reminder that the path ahead may be arduous and lengthy; however, those who exhibit persistence and determination will be abundantly rewarded. For instance,

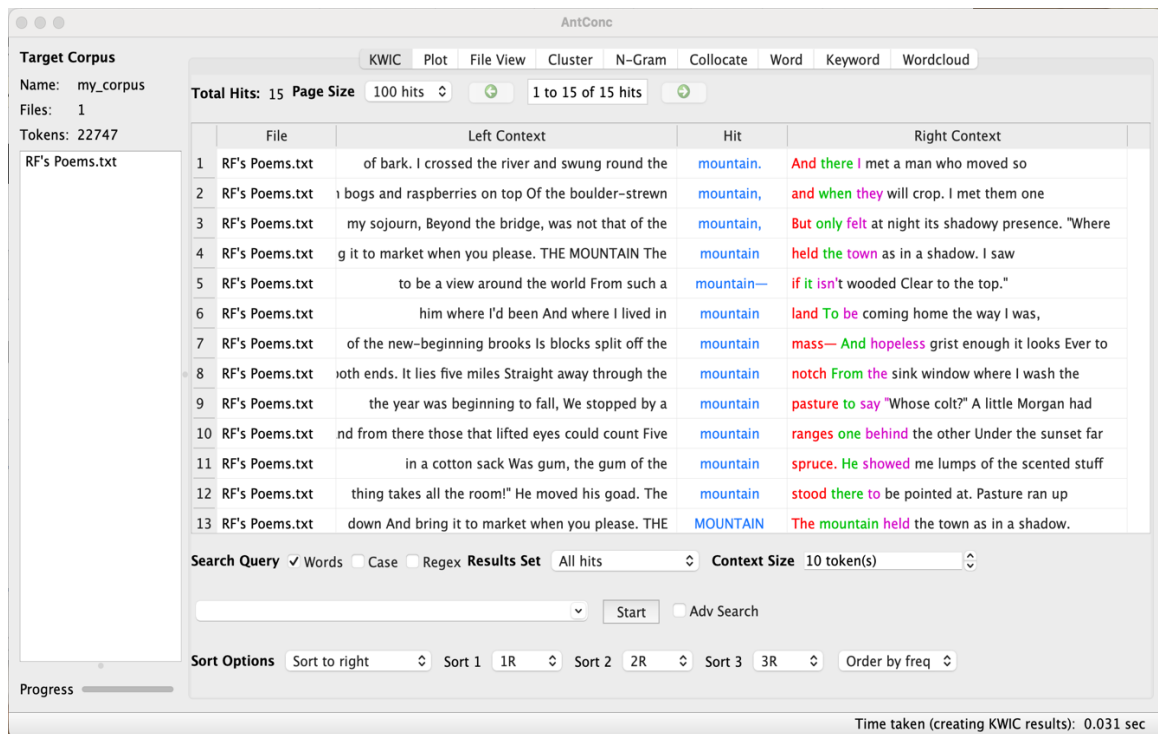


FIGURE 5. KWIC (Key Word in Context) of the word, "mountain"

According to Figure 5, RF's perspective on how the mountain is a core focus of the poem encourages people to think, to be determined, and to be curious because it is depicted as an extensive, unmoving force. For lack of a better explanatory term, it would be said that mountain is used in the poems but not in the particular core sense of the word. It is turned into a figurative

term instead, which helps to accentuate the desolated nature of the artwork. In addition, while the KWIC of the word "mountain" does not directly represent a mountain in the text, the ideas of elevation and climbing seem to persist throughout. The earth retains some features of a mountaintop, reflecting the willingness to strive for more. Such a notion of climbing or raising one hand to touch the firmament captures the pictorial idea of a mountain, even if it is visualized differently.

Therefore, it is vital to tackle the relationship between mountains and human experience. Mountains can represent the pinnacle of human aspiration and achievement. When mountains are conquered, it is proof to us and others about our capability of greatness. Mountains can also be a symbol of the unknown and the mysterious. They can represent the fear and uncertainty we feel when faced with the unknown. Finally, mountains can represent the beauty and wonder of the natural world. They are a reminder of the power and majesty of nature and the importance of preserving it for future generations. The utilization of mountain imagery by RF serves as a manifestation of his inclination for traditional American values and his skepticism towards modernity (Elder, 1998). Lawrence Buell (1998) commented that mountains in RF's poetry represent a response to the historical and cultural contexts of his time, including his interest in Eastern philosophy and nature writing. For RF, mountains are a symbol of the human struggle to achieve self-realization and spiritual growth. Even though RF wrote his poetry over a century ago, his use of mountain imagery continues to resonate with readers today, reminding us of the enduring power of nature and the importance of perseverance in achieving our goals.

**Barn:** Barn was mentioned 28 times. Barns were essential for storing crops, livestock, and farming equipment, with distinct architecture and communal activities. They symbolize the human desire to create lasting, meaningful things in a world of impermanence. Barn imagery explores themes of aging, decay, and loss, capturing the power, fragility, and mystery of rural life and nature's grandeur. "That's what brings me to barns," wrote RF in a letter to Louis Untermeyer on March 5, 1921. Barns represent the human desire to create something lasting and meaningful in a world that is characterized by impermanence and change. RF uses barns as symbols of a disappearing way of life in rural areas, highlighting the tension between tradition and progress. The poems "Mending Wall" and "The Death of the Hired Man" explore themes of aging, loss, and contradictions, while "Birches" reflects the transience of life and the human search for meaning in a chaotic world.

**Wall:** The word 'wall' has been mentioned 23 times in the corpus. The 'wall' in the poem stands for the duality of life, the subject of destruction and creation, which are interdependent. Here, it is noticeable that destroying ability does not always indicate negativity if it inhibits decent neighbors from getting along, just as creating something that alienates people from one another is not always a good thing, such as using the wall between neighbors can be enough. The RF's wall, which simultaneously divides and brings two individuals together, is a representation of ambiguity (Srinivasarao, 2013). The poem "Mending Wall" by RF explores the metaphorical sense of walls dividing people, serving both positive and negative purposes. It depicts estranged men gathering annually to rebuild a wall made of boulders, imagining it as shared loaves and toys.

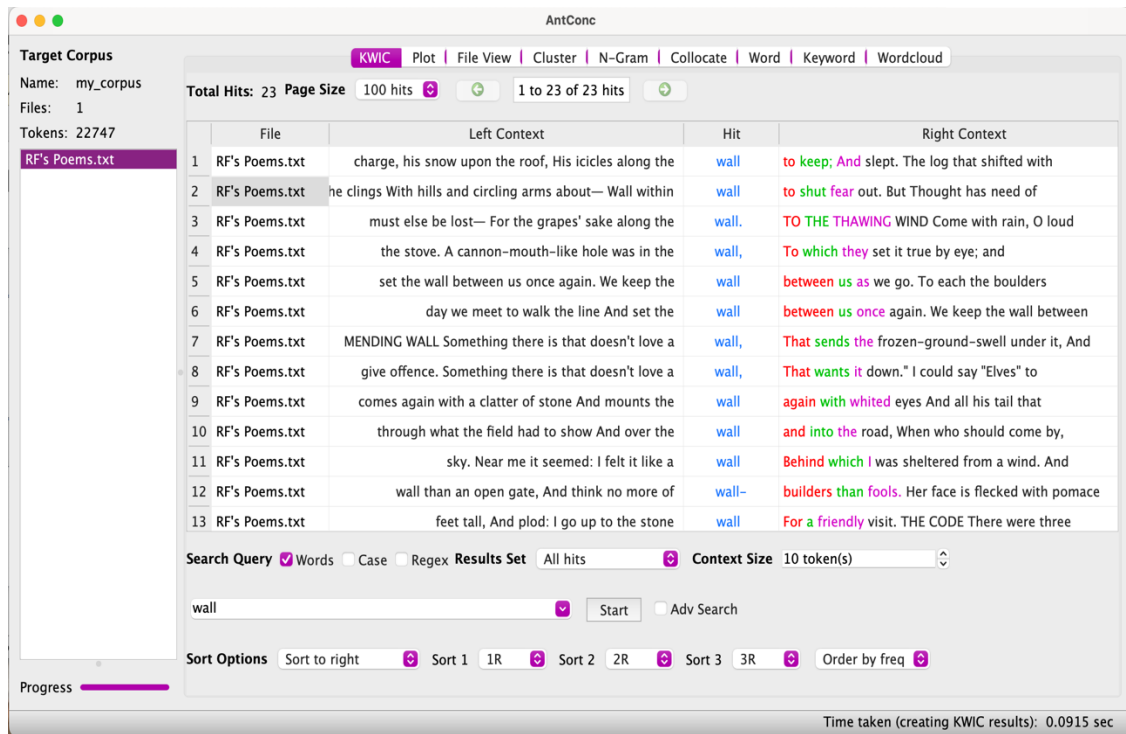


FIGURE 6. KWIC (Key Word in Context) of the word, “wall”

**Way:** The word ‘way’ appeared 124 times in the node corpus, which indicates its high impact and value in RF’s poems. The “way” is a recurring theme in RF’s poetry. It can refer to a path, a journey, or a decision that one must make in life. The “way” is often used as a metaphor for the choices we make in life and the paths we take. RF’s poems suggest that the choices we make in life can have a profound impact on our future.

The poem “The Road Not Taken” explores themes of choice, uncertainty, indecision, fate, and overthinking. The speaker is uncertain about the best path and wonders if it’s suitable for them. The word “way” serves as a metaphor, referring to the path, road, and two roads, and comparing it to other themes in RF’s poetry provides a deeper understanding.

Another example is “Stopping by Woods on a Snowy Evening.” The “way” in this poem refers to the speaker’s journey towards death. The woods represent death, and the “way” is the path he must take. The ending of the poem suggests that the speaker is not ready to take the final “way” yet. “The Death of the Hired Man” explores the theme of loyalty and the choices we make in life. Warren must choose between loyalty to his wife and loyalty to his friend, the hired man. The “way” in this poem represents the two paths that Warren must choose between.

**Snow:** Snow has been mentioned 42 times. In general, snow signifies winter. It deals with the setting of the poem, which takes time in winter season. It indicates the end of the growing seasons. Because of its relation to the end of the growth in the natural world, snow is often linked to the death and hardship. For example, early Americans gathered food during spring, summer, and fall to use during winter months. Anyone who failed to gather these crops during the growing season experienced difficulties during the winter months. On the other hand, eastern cultures also use the color white to represent mourning and death (Afrianto & Restika, 2018; Kuswanto et al., 2021; Renaldi et al., 2016).

The writer suggests that snow symbolizes death, hardship, and bleakness, relating to the first symbol representing trouble and despair. These hard conditions exist to measure one's strength, as seen in the first and second lines of the "Storm Fear" poem.

RF's poems often use snow imagery to convey deep human emotions, such as alienation from society, loss of innocence, and tree damage. His mastery of using nature and the outdoors to express emotions is evident in his use of snow in poems like "The Hill Wife," "Directive," and "Birches." In addition, RF's poetry often uses snow in particularly unique ways. It's not limited to just a symbol of death, but rather the different interpretations based on context that make it so rich. The complex nature of the element itself allows him to convey various themes such as loss of innocence, clarity of mind, nature's beauty and cruelty, our relationship with the unknown, and the vastness of existence. RF's poetry uses snow to symbolize coldness, death, loneliness, and contemplation, elucidating central themes. His masterful stylization of snow highlights his contributions to American poetry, capturing universal themes and the enduring nature of poetry.

## CONCLUSION

The present study demonstrated that RF's poems' themes can be interpreted based on the frequency of use of keywords, through employing corpus linguistics methods. The most frequent keywords that function as indicators of the nature theme in RF's selected poems are mountain (15), barn (28), wall (23), way (124), and snow (42). Therefore, the researcher concluded that the registered keywords in RF's poem texts in English confirm how the natural elements are used as a background for the other themes in his poems because the keywords mentioned above are used in most of the texts of RF's poems (Zhang et al., 2017).

RF, a poet of nature, explores the interplay between man and nature in his landscapes. He highlights the complex relationship between humans and nature, with humans both protecting and destroying it. RF's poetry often explores the contrasting feelings of kindness and evil towards nature, with some people fearing its tyranny and others bravely resisting it. Some people view nature as a source of revelation, while others view it as a threat. RF's work highlights the complex relationship between humans and nature. RF's work often begins with an observation of nature and then connects it to psychological concerns. He sees nature as a representation of humanity's connection to the external world, exploring elements like forests and wildflowers. RF's work reflects the profound conundrum of human existence, highlighting the ongoing conflict between imagination and existential emptiness.

RF was often inspired by the natural world around him. He used nature to explore complex themes such as mortality, human relationships, and the meaning of life. RF's poems often contrast the beauty and harmony of nature with the chaos and violence of human society. He believed that nature had a healing power that could help us to understand ourselves and the world around us.

In conclusion, the exploration of the theme of nature in RF's selected poems is possible through the use of corpus linguistics, by analyzing keywords, collocations, and colligations in the corpus. AntConc helped in bringing out the unusually frequent words related to the theme of nature in the poems. The three questions were answered via the use of the corpus linguistic approach. Nature was represented in the form of the natural elements selected by Robert RF and constituted RF's lexical choice. Keywords, collocations, and colligations contributed to the theme of nature and made the readers understand the reason for using pastoral and natural elements in his poems. The corpus stylistic approach has been proven to be a successful method to provide us with a better



understanding of the poems. While reading RF's poetry, it is apparent to see the influence of New England on his writing. Whether this influence was positive or negative, his poems have been read by millions over the years and have grown in popularity year after year (Buxton, 2004).

## REFERENCES

- Abel, D. (1981). Unfriendly nature in the poetry of Robert Frost. *Colby Quarterly*, 17(4), 4.
- Adams, J. D. (1966). *Speaking of books and life*. Holt, Rinehart and Winston.
- AlKhrisheh, A. S., Abuhammam, E. A., & Al-Rbehat, W. R. (2023). Questioning nature: A study of death and isolation in selected nature poems by Robert Frost. *Theory and Practice in Language Studies*, 13(1), 127–136.
- Almiqdady, M. Z., Abu-Melhim, A. R. H., & Al-Sobh, M. A. (2015). The Concept of nature in the poetry of William Wordsworth and Robert Frost: A comparative study. *Arab World English Journal (AWEJ) Special Issue on Literature*, (3).
- Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, 9(3), 212–230.
- Baym, N. (1965). An approach to Robert Frost's nature poetry. *American Quarterly*, 17(4), 713–723.
- Brezina, V., Weill-Tessier, P., & McEnery, A. (2021). #LancsBox v.6.x (Version 0.6) [LancsBox]. [<http://corpora.lancs.ac.uk/lancsbox/download.php>].
- Buxton, R. (2004). Robert Frost and Northern Irish poetry. OUP Oxford.
- Coletta, W. J., & Tamres, D. H. (1992). Robert Frost and the poetry of physics. *The Physics Teacher*, 30(6), 360–365.
- Culpeper, J. (2001). *Language and characterisation: People in plays and other texts*. Pearson Education.
- Dabirnia, M. (2021). A comparative study of the representation human/nature relationship in selected poems from Robert Frost and Sohrab Sepehri from an ecocriticism perspective. *Advances in Language and Literary Studies*, 12(6), 87–94.
- Dougherty, J. P. (1966). Robert Frost's "Directive" to the wilderness. *American Quarterly*, 18(2), 208–219.
- Elder, J. (1998). *Reading the mountains of home*. Harvard University Press.
- Fagan, D. J. (2007). *Critical companion to Robert Frost: A literary reference to his life and work*. Infobase Publishing.
- Faggen, R. (1997). Robert Frost and the challenge of Darwin. University of Michigan Press.
- Faggen, R. (2014). Robert Frost. In D. E. Friedman & C. G. Nally (Eds.), *A companion to modernist poetry* (pp. 358–366). Wiley-Blackwell.
- Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In *Studies in linguistic analysis* (pp. 1–32). The Philological Society.
- Fischer-Starcke, B. (2010). *Corpus linguistics in literary analysis*. Continuum.
- Frost, R. (1978). *Stopping by woods on a snowy evening*. Irwin Limited.
- Hass, R. B. (2005). (Re)reading Bergson: Frost, Pound, and the legacy of modern poetry. *Journal of Modern Literature*, 29(1), 55–75.
- Hota, S., Argamon, S., Koppel, M., & Zigdon, I. (2006, July). Performing gender: Automatic stylistic analysis of Shakespeare's characters. In *Proceedings of Digital Humanities 2006* (pp. 100–104).

- Jaafar, E. A., & Ganapathy, M. (2022). Investigating EFL learners' ability to analyse poetic language: A pedagogical corpus stylistic approach. *Theory and Practice in Language Studies*, 12(5), 866–875.
- Kemp, J. C. (2015). *Robert Frost and New England: The poet as regionalist* (Vol. 1430). Princeton University Press.
- Kilcup, K. L. (1998). *Robert Frost and feminine literary tradition*. University of Michigan Press.
- Klein, A. (2008). The counterlove of Robert Frost. *Twentieth Century Literature*, 54(3), 362–387.
- Kuttapan, J., & Joy, G. (2023). Analysis of the obscure biblical allusions in Robert Frost's "Mending Wall." In *9th International Scholars' Conference Proceedings*.
- Lynen, J. F. (1960). *Nature and pastoralism*. Yale University Press.
- Mahlberg, M. (2013). *Corpus stylistics and Dickens's fiction*. Routledge.
- McIntyre, D. (2013). The year's work in stylistics 2012. *Language and Literature*, 22(4), 333–347.
- Monteiro, G. (1988). *Robert Frost and the New England Renaissance*. University Press of Kentucky.
- Montoro, R. (2015). Review of *Digital literary studies: Corpus approaches to poetry, prose, and drama*, by D. L. Hoover, J. Culpeper, & K. O' Halloran. *International Journal of Corpus Linguistics*, 20(1), 129–137.
- Murphy, A. (2006). The Oxford Shakespeare re-viewed (again). *Text*, 16, 157–171.
- Naess, A. (2015). The deep ecological movement: Some philosophical aspects. In K. Hiltner (Ed.), *Ecocriticism: The essential reader* (pp. 47–61). Routledge.
- Rashid, M. H., Husain, H. J., Islam, J., Sipahi, E., & Hui, W. (2021). Explore the natural beauty of Robert Frost in his poetry. *International Journal of Applied Linguistics and English Literature*, 10(3), 15–19.
- Selvi, M. T., & Malar, A. A. (2021). A critical analysis of Robert Frost poems. *Journal of Language and Linguistics in Society*, 1(2), 1–6.
- Srinivasarao, S. (2013). Nature poetry and modernity: A critical review on Robert Frost poetry. *Research Journal of English Language and Literature*, 1(1), 161–164.
- Stubbs, M. (2002). On text and corpus analysis: A reply to Borsley and Ingham. *Lingua*, 112(1), 7–11.
- Thakur, S. K. (2022). Robert Frost's vision of life. *Ideal Research Review*, 70(1), 58–62.
- Thompson, L. (1966). *Selected letters of Robert Frost*. Holt, Rinehart and Winston.
- Toolan, M. (2009). *Narrative progression in the short story*. John Benjamins.
- Verdonk, P. (2013). *The stylistics of poetry*. Wiley-Blackwell.
- Zeng, Y. (2024). Brief analysis of the characteristics of Robert Frost's poetry. *International Journal of Education and Humanities*.
- Zhang, Y., Ding, W., & Jia, L. (2017). Analysis on nature in Robert Frost's poetry. *English Language, Literature & Culture*, 2(3), 25–30.
- Zyngier, S. (1999). "Smudges on the canvas"? A corpus stylistics approach to *Macbeth*. In I. Biermann & A. Combrink (Eds.), *Poetics, linguistics and history: Discourses of war and conflict* (pp. 529–545). Potchefstroom University.

### **ABOUT THE AUTHORS**

Amira Hassan Ahmed Abdullah Saad is a researcher specializing in corpus stylistics. A former language instructor at Jazan University, she now works at Jazan Institute and is a PhD candidate at UNIMAS, Malaysia. Her research explores linguistic patterns in texts, using computational tools to analyze style, meaning, and authorial voice.

Monaliza Sarbini-Zin (Ph.D), a senior lecturer at UNIMAS, specializes in linguistics, critical stylistics, and discourse analysis. She co-authored a 2021 study on ideology in Kipling's "If". Actively involved in teacher training and curriculum development, she collaborates with Malaysia's Ministry of Education to enhance teaching standards and linguistic studies.