



Title	Between Distant and Close : Sea and Inland in Tennyson by Topic Modelling
Author(s)	藤田, 郁
Citation	言語文化共同研究プロジェクト. 2025, 2024, p. 19-41
Version Type	VoR
URL	https://doi.org/10.18910/102220
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Between Distant and Close: Sea and Inland in Tennyson by Topic Modelling

Iku Fujita

Graduate School and Faculty of Information Science and Electrical Engineering,
Kyushu University

744 Motooka, Nishi-ward, Fukuoka city, Fukuoka, 819-0395 Japan

E-mail: tnnysn.annie.if@gmail.com

Abstract There have been many earlier studies on Alfred, Lord Tennyson, a representative poet of nineteenth-century English literature. Despite the fact that quantitative analyses have been used with prose texts, studies of Tennyson and other poets have rarely used such numerical approaches meanwhile they have employed the qualitative approach based on their close reading. This study reads Tennyson's poems from a distance by using a topic model, a quantitative technique. The analyses are based on the results of the latent Dirichlet allocation topic model (LDA), and examine how the poet depicts his themes by using imagery of the sea and inland environments, the latter including mountains and rivers. In the emerging results of LDA, the sea tends to convey the impressions of despair and eternity (after life) with the image of the broad sea itself, while inland landscapes of mountains connote a religious sense and rivers are used figuratively as life, but not hopelessness or perpetuity. The study's findings are not necessarily novel, but they do demonstrate how integrating quantitative and qualitative methodologies can be beneficial. Despite these components, neither the author nor the topic modelling stressed that a poem should only have one motif, theme or topic. The researcher 'reads' Tennyson's poems from a distance and up close while weighing the advantages and disadvantages of qualitative and quantitative approaches.

Keywords Alfred Tennyson, distant reading, LDA, poetry, topic model

トピックモデリングを用いた テニスンにおける海と陸の考察

藤田 郁

九州大学大学院システム情報科学研究院

〒 819-0395 福岡県福岡市西区元岡 744 番地

あらまし 本研究は、定量的手法であるトピックモデルを用い、19世紀イギリス文学を代表する詩人の一人、アルフレッド・テニスンの詩を分析、遠読するものである。テニスンやテニスンの作品に関する先行研究は数多く存在する。しかし、テニスンをはじめ、他の詩人や彼らの作品に関する先行研究では、定量的なアプローチが用いられることはあまり多くなく、精読に基づく定性的なアプローチが主として用いられてきた。本研究では、トピックモデルの中でも潜在ディリクレ配分トピックモデル（LDA）を用い、その結果に基づき、テニスンが作品内でどのように海と

内陸環境（山や川）のイメージを用い、テーマを描いているかを考察する。これまでの研究では、テニスの作品における海と死の関連が指摘されていたが、LDAの結果、海は死だけでなく、広い海のイメージから、絶望や永遠性（死後の世界）の印象を伝える傾向があることが示唆された。また、内陸の風景である山は宗教的な意味を、川は人生を比喩的に表すために使われており、海に見られるような絶望や永遠といったイメージは含まれていないと解釈される。本研究によって得られた結果は、例えば海と死の関連など、必ずしも目新しいものではない。しかし、定量的手法と質的手法を融合することで、これまでの研究で指摘されてきたテニスの作品におけるテーマや情景の描写が、特定あるいは単一の作品のみではなく、複数の作品に跨がっていることを明らかにしたほか、海や川といった要素が特定の概念だけではなく、様々な感情を含意、増幅する効果があることが明らかになった。

キーワード アルフレッド・テニス、遠読、LDA、詩、トピックモデル

1. Introduction

This paper is a part of exploratory analysis and study of the verse texts of the nineteenth-century Victorian poet, Alfred, Lord Tennyson (1809–1892). The analysis of this study employs an approach of stylometry, also called computational stylistics, ‘is concerned with the quantitative study of writing style’ (Eder et al., 2016: 107). Numerous studies have examined Tennyson’s works by focusing on similarities and differences between Tennyson’s and other poets’ styles, syntax, or lexica. However, few studies have employed quantitative methods, whereas most studies take a qualitative approach to specific poems based on close reading.

Moretti (2013) indicates that the problem with qualitative literary studies is that they inevitably only rely on a few works or expressions of a writer. Furthermore, they only focus on a limited portion of a particular work by a writer, thus omitting the remaining large portion of the work. Moreover, it is not necessarily true that an exhaustive analysis is possible even if one does take the time to do so. In fact, more than 130 years after the publication of Tennyson’s last work, there has still been no macroscopic or comprehensive study of Tennyson’s work as a whole or a comparison of his works with those of other authors using qualitative methods. The term ‘distant reading’ was referred to by Moretti and is nowadays often used in contrast to ‘close reading’. Distant reading enables us ‘to focus on units that are much smaller or much larger than the text: devices, themes and tropes’ (Moretti, 2013: 48–49).

This paper employs the latent Dirichlet allocation topic model (Blei et al., 2003; hereafter LDA) as a stylometric approach. LDA is a method specifically designed for analysing large text data. Whereas previous studies that have employed LDA have mainly focused on prose texts, this study attempts to apply LDA to poetry data and obtain efficient results and insights. LDA enables distant reading by identifying underlying semantic relationships between words using probabilistic calculations. This study not only demonstrates the usefulness of the method but also presents new features produced by the method, thus reinforcing previous discussions. This study expands on the previous study (Fujita, 2023; see details in Section 2.2), which presented comprehensive results

using the same methodology as this study, but further focused on specific topics and broke down the findings. The slight alteration of the data set from Fujita (2023) has been made; however, the altered data set does not negate or contradict the findings of both this study and Fujita (2023).

It has been noted by several scholars that Tennyson had a particular affinity for bodies of water, such as the sea and the lake. However, comparatively little attention has been given to rivers in Tennyson's oeuvre by scholars. Instead, it is the case that rivers in Tennysonian works attracted readers, and 'The Lady of Shalott', for example, has a notable presence and has been depicted in paintings and other visual arts. It is evident that scholars and readers' notions, as well as the following and Fujita's (2023) LDA results, have identified and classified the sea and river as discrete categories (topics). This study will therefore focus on the portrayal of the sea and the inland, including rivers, in Tennyson's works.

In addition to the stylometric approach, this study also employs qualitative methods in its discussion. Given the emerging results of the quantitative method, LDA, this paper conducts a close reading of Tennyson's poems, along with previous studies. By combining distant and close readings of Tennyson's poems, this study sheds light on new aspects of Tennyson's works. Computer analyses can provide clues to draw human attention to unintentionally ignored aspects of poetry, however, it is the responsibility of individuals to read, interpret, and engage in discussions about poetry. Neither computers nor humans are perfect. The author of this study therefore expects that quantitative and qualitative methods can be complementary in order to extend our knowledge and to identify new and unperceived aspects of literature. This study also aims to demonstrate that stylometric approaches and 'distant reading' can expand the scope of analysis and provide a bird's-eye view.

2. Literature Review

2.1. Previous studies on Tennyson

There are abundant previous studies of Tennyson's life and his poetry, style and prosody. *Tennyson: A Memoir* (H. Tennyson, 1897) and *Alfred Tennyson* (C. Tennyson, 1949) are undoubtedly the closest source biographies to Tennyson, having been written by Alfred's son, Hallam, and his grandson, Charles. Brooke (1894) is nominated an earlier comprehensive study on Tennyson's poetry that was published only two years after Tennyson's death. In addition, Dixon's (1896) survey ranges from *Poems by Two Brothers* (1827) to *Ballads and Other Poems* (1880) and included a detailed biography of Tennyson.

Ricks (1969, 1987a, b, c) never allowed other authors to overtake him in Tennyson annotations. Indeed, previous researchers such as Shaw (1973) and Hair (1991), and later studies including Thomas (2019) refer to Ricks' both works. Ricks identifies references to other authors in Tennyson's poems, which Shaw (1973) confirms later. Shaw also counts the number of such references and expands his own discussion of Tennyson's style. Hair (1991) identifies qualitative

features of Tennyson's poetry, such as its vocabulary, concepts or motifs, by comparing them with those of other poets. In a later critical study, Thomas (2019) highlights 'echoes' of Wordsworth's poems in Tennyson's works.

H. Tennyson (1897: 20) notes that from his childhood, Alfred Tennyson 'had a passion for the sea, and especially of the North Sea in wild weather', as well as emotional reactions to water: "‘‘Somehow,’’ [Tennyson] would say, “water is the element I love best of all the four”" (H. Tennyson, 1897: 49). Further, Okazawa (1969) explains Tennyson's love for, enthusiasm for, and exploration of waters such as seas, lakes, and rivers. Readers can find both seas and lakes in more than 20 poems, including 'Morte d'Arthur', 'Edwin Morris, or The Lake', and 'The Ring' and rivers in over 40 such as *The Princess*, 'The Lady of Shalott', and 'The Brook'.

For years, Tennysonian scholars have found the sea to be a more intriguing motif than rivers as well as a common motif in his poems. Keirstead (2019) broadly explains how Tennyson evokes the sea, citing works such as 'Mariana in the South', 'Enone', 'The Lady of Shalott', 'Locksley Hall', 'Ulysses', *In Memoriam A.H.H.*, *Enoch Arden* and the *Idylls of the King* series. He describes 'Simply put, going to sea means almost certain death in Tennyson' (Keirstead, 2019: 75). However, he also suggests that when a beach or seashore is depicted in a poem, Tennyson is also expressing or connoting aspects such as comfort or unstableness that readers can receive from beach/seashore images (Keirstead, 2019: 84).

Keirstead is the most recent scholar to identify the sea motif in Tennyson and analyses a broad swath of his work with a qualitative approach. Besides, Fulweiler (1965) also examines what sea portrays, connotes, and relates to in Tennyson's poetry; he discusses, for instance, 'The Merman', 'The Mermaid', *In Memoriam A.H.H.*, 'Ulysses', 'Crossing the Bar', 'The Holy Grail' and 'Morte d'Arthur'. Although the sea where the dying Arthur heads for in 'Morte d'Arthur' and its later version, 'The Passing of Arthur', is a huge lake, Tennyson describes it with the words *sea*, *ocean*, *water* and *lake* in the two poems. Thus, Keirstead (2019) and Fulweiler (1965) discuss the lake in 'Morte d'Arthur' and 'The Passing of Arthur' as an inland sea. These previous studies are qualitative, which is a fairly standard approach in poetry research.

Whereas many of these researchers take a qualitative approach to studying Tennyson's poems, Plamondon (2005) is one of the few who employ a quantitative approach; he investigates differences in the frequency of some vowels and consonants between Tennyson's and Robert Browning's poems.

One of Tennyson's characteristics is his long-form poetry. He wrote over 400 lyrical poems, with word tokens ranging from 12 to 300, but critics can never pass over Tennyson's narrative and epic poems, which are often longer than the lyrical poems. For example, his longest poem is *The Princess*, whose word-token count is 26,526. Furthermore, his most representative poem, *In Memoriam A.H.H.*, is a lyrical poem that is also his second-longest work with 18,724-word tokens. Although Tennyson has more short lyrical works than long or narrative poems, these eye-catching

long poems are never neglected and are repeatedly observed by critics. Even though his works contain shared elements beyond individual poems, qualitative researchers can sometimes miss these elements when they focus on only representative but small subsets of Tennyson's works.

2.2. Stylometric approaches and topic model

The qualitative approach is intimately associated with the practice of close reading. This method is considered indispensable and consequential in literary research. Nevertheless, Moretti (2013: 48) states that 'the trouble with close reading [...] is that it necessarily depends on an extremely small canon'. He further explains that given that close reading focuses on a limited portion of the work, it leaves 'great unread' aspects outside of 'an extremely small canon'. Here, why he employs the phrase 'extremely small' is because the quantitative can handle a relatively larger amount of data at once while the qualitative finds it challenging to handle not only a small canon but non-canonical works as well.

'Distant reading' utilises quantitative approaches and can be effective in exploring beyond the canon, as Moretti suggests (2013: 48). However, neither I nor Moretti (2013) intend to criticise research based on close reading or the approach itself. As few researchers use statistical and stylometric approaches, this paper aims to identify the 'great unread' in Tennyson's poetry employing a quantitative method called LDA (Blei et al., 2003).

The basic idea of LDA is that 'documents are represented as random mixtures about potential topics, where each topic is characterised by a distribution over words' (Blei et al., 2003: 996). Briefly, LDA classifies words into groups called topics. The algorithm and emerging results reveal 'the semantic structure hidden in a corpus of text' (Tabata, 2018: 52). Additionally, LDA is a specialised method for analysing big text data (Iwata, 2015).

Several researchers have employed topic models to examine prose texts (e.g., Onodera et al., 2016; Kuroda, 2017; Tabata, 2017, 2018, 2020; Kiyama, 2018; Matsukawa et al., 2018; Huang, 2020a, 2020b). However, its applications for poetry are fewer than those for prose text studies (e.g., Rhody, 2012; Navarro-Colorado, 2018; Henrichs, 2019; Okabe, 2019). Although this can be seen as a consequence of difficulty in applying LDA to poetry, Fujita (2022) has shown that LDA is effective when examining Tennyson's poetical works. While Fujita (2022) focuses on the methodological conditions of applying LDA to Tennysonian works, Fujita (2023) claims the general results of exploratory analysis by LDA on Tennyson's poetry. Fujita's (2023) findings include the different depiction of sea and river as well as the association between the environmental materials and the emotion of characters, and the findings are closely related to this article, but it is more general suggestions than what LDA detected than that giving deep discussion on the themes. Thus, this study aims to gain more specific discussion on the specific themes that LDA found including more number of quotations and mentions on the poems themselves. With this earlier study as an empirical ground for justifying the use of LDA, I analyse 593 works of Tennyson, focusing particularly on the nouns he uses in this paper. The next section will show more detailed information about the data set of 593 poems, and about the methodology for the analysis.

3. Methods and materials

3.1. Data sources and configuration

This study is an analysis of 593 poetic works written by Tennyson. Of the 593 poems, 423 were compiled from the *Delphi Poets Series* ‘Alfred, Lord Tennyson’ (2013), and the rest, 170, were from Ricks (1987a, b, c). All the texts were converted to plain text data by optical character read/recognition, and then it was manually proofread and edited for errors such as unnecessary spaces, tabs, newlines and garbled characters. Table 1 shows the statistical data of the Tennyson corpus, 593 digitalised poems.

Table 1: Statistical data of Tennyson corpus

The number of poems	593
Total tokens	351,063
The shortest poem in a number of words	12
The longest poem in a number of words	26,383
Mean tokens per poem	592.01
Standard deviation	1825.61

The texts of the works were given part-of-speech (PoS) tags using CasualConc 2.0.8 (Imao, 2022)¹. The tagset is CLAWS5, the same tagset as used in the British National Corpus. To analyse components, which heavily reflect themes, topics and subjects of poetry, this analysis adopted only nouns that were suggested by the PoS tags as ‘_NN1’ (singular nouns), ‘_NN2’ (plural nouns) and ‘_NN0’ (mass nouns). All other words were deleted before the LDA application after the works were separated into consecutive segments (See Section 3.2 for more details about segments).

3.2. Modelling and data analysis methods

LDA calculates the probabilities of word frequencies and classifies words into groups, called topics, on the basis of the hypothesis that documents should be classified into several groups based on their tendency to co-occur (within the same document). Here, documents are defined as contiguous segments of textual data, and they are not necessarily equivalent to poetical or prose work units. Fig. 1 shows the sample results of running LDA (Blei et al., 2003: 1009). LDA analysed the documents in the lower part of the figure and classified words into topics. LDA outputs topics and other numerical results such as word-weight values for each word, alpha (α) values for topics and the density of topics in work, but did not produce the labels, ‘Arts’, ‘Budgets’, ‘Children’ and ‘Education’ in Fig. 1. The analysts of the study interpreted LDA results by combining their knowledge of the text data and assigning labels to each topic.

LDA employs a bag of words model, which ‘treats each document as a bag of words’ and ignores the order of words in documents (Jockers, 2014: 137). Regarding his own novel text

¹ Notwithstanding the perpetual updating both of CasualConc itself and its manual, it should be noted that this article is referring to the version of the program that was available at the time this study was conducted.

“Arts”	“Budgets”	“Children”	“Education”
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

The William Randolph Hearst Foundation will give \$1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. “Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services,” Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center’s share will be \$200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive \$400,000 each. The Juilliard School, where music and the performing arts are taught, will get \$250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual \$100,000 donation, too.

Figure 8: An example article from the AP corpus. Each color codes a different factor from which the word is putatively generated.

Fig. 1: Example text from the AP Corpus (cited by Blei et al., 2003: 1009).

analysis, Jockers (2014) explains that LDA captures ‘some themes that run throughout and others that appear at specific points and then disappear in novels’, and ‘it is useful to divide novels (and other large documents) into chunks or segments and then run the model’. Further, LDA requires a significant number of documents to make its probability calculation sufficiently accurate.

In previous LDA studies on prose texts, works were split into segments of equal sizes, such as 1,000 (Kuroda, 2017; Tabata, 2020), 2,000, or more words (Huang, 2020b). On the other hand, poetry texts in studies employing LDA have not been sliced into segments; rather, scholars have considered poems themselves as segments. Navarro-Colorado (2018) and Henrichs (2019) apply LDA only to sonnets, whose tokens range from about 100 to 200. Meanwhile, Tennyson’s poems vary in length, as the standard deviations in Table 1 show. The large deviation of word tokens interferes with LDA results because a large population difference can affect the reliability of LDA calculations and results.

Further, the large disparities in segment size, instantly exert an influence on the results because LDA uses raw word frequency to identify the co-occurrence of words. Therefore, it is advisable to divide Tennyson’s poems into segments of possible equal size following the method of prose text studies considering the large deviation of tokens per poem. Jockers (2014: 137) additionally mentions the size of segments: ‘the bigger the bag, the more words that will tend to be found together in the same bag’. Nonetheless, careful segment sizing is necessary because in the reverse

of Jocker's words, the smaller the bag, the fewer the words that will tend to be found together in the same bag.

Segment size should be set up for appropriate chunks that contain coherent topics in poems. The results will not fruitfully reflect the essences of the works when the poems are divided into too small segment size or concatenated individual poems. Therefore, segment size should be set to minimise differences between segments. Regarding these concerns, Fujita (2022) suggests the relevant segment size for LDA on Tennyson's poems. This paper refers to Fujita and adopts a segment size of 592 words based on the mean number of tokens in the Tennyson corpus shown in Table 1.

After that poems were split into 592-word segment size, all words besides nouns (see Section 3.1) were deleted. LDA was applied to the segments using the Machine Learning for Language Toolkit (McCallum, 2002). The number of topics was set at 18 considering results of experimental trials, with the number of topics ranging from 10 to 200.

4. Results

This section presents the results of LDA for the Tennyson corpus. Table 2 shows one set of output results, the α value for the 18 topics, the 20 most salient keywords for each topic and the labels, which the author of this paper assigned manually. The keywords are ranked in descending order of their weights from the upper left to the bottom right of the table; α value represents the universality of each topic. The higher the value, the more prominently the topic appears in multiple works; lower values indicate that a topic appears in fewer works, sometimes only one. Topic 1 in Table 2 had the highest α , 0.8774, of the 18 topics, indicating its greater prominence than all the others; Topic 8 was the least prominent ($\alpha = 0.0581$). Detailed analysis of the poems and the LDA output results suggested specific labelling for each topic. The labels show that both Topics 2 and 3 refer to natural environment. Specifically, Topic 2 pertains to seas and their surroundings, while Topic 3 pertains to natural ambience inland. The top 20 keywords for both topics include words related to bodies of water, seas, and rivers. Although previous studies have discussed seas in Tennyson's work, rivers have been scarcely mentioned. The α values of Topics 2 and 3 are similar, indicating that both topics are present in the corpus to a comparable extent.

Table 2: Output results of LDA: 18 topics with their labels, alpha values and keywords

Topics	Alpha Values	Labels	Keywords (1–20)
0	0.2177	Nobles	king queen knight knights hall man maid name field court face table day quest brother men realm hands sword prince
1	0.8774	Life (materials)	light sun day eyes heaven time voice earth shadow wind moon sea cloud sound stars hills land morn glory hill
2	0.1345	Environment (sea)	sea isle shore day boat sail seas ship wave side ocean home bells sand tide blast sails vessel storm beach
3	0.1266	Environment (in-land)	mountain river valley melody music city palm bridge pine mould tree prime water lake garden gleam fountain torrent song level
4	0.6847	Male	man life world things time men power years truth mind nature days voice friend heart peace earth change faith hands
5	0.2289	Natural scenery	year rose flower flowers summer garden birds roses air brook leaves feet tree spring leaf morning woodland day blossom wood
6	0.6871	Life (emotion)	love heart eyes life death tears hand soul face name lips beauty light cheek hope youth side dream blood smile
7	0.0592	Ladies	lady princess ringlet woman prince honour college goose gown head girls girl sister highness women letters echo web mirror books
8	0.0581	Discretions in life with seasons and nature	fool mark wine men leaf tongue cup swine ways hair innocence moment rhymes thing port lodge lies pity summers cups
9	0.0823	Attractiveness	eye air sense charms bloom face maze harmony lustre language depth horizon dye course ecstasy spell mysteries flash claim shrine
10	0.2844	Man and death	death man fire war god blood men hell gods way heaven hand earth battle head flesh foe flame king shame
11	0.0903	Woman and whereabouts	woman dream gold boon child men charm faith eye music belt books ruin town wrongs mood cave wife heaven street
12	0.0920	Knighthood	horse hall prince arms wood damsel knight shield lady knave pride walls horses squire bridge women armour charger lance town
13	0.3860	Family	mother child father heart man hand wife face day house boy men head bride woman word children will home daughter
14	0.1400	Battle and belief	land name men fame people war freedom hearts cause crowd sons friends kings throne queen shame peace health tyrant greatness
15	0.0505	Cheering finiteness	ring laughter t'amo ward finger bone harp bells month spring lisette cousin souls babble noise smile chest spirits landscape marge
16	0.0370	Religions and life outside the UK	children bread people chains church enemy moor heresy women colony chorus roof east banner temple goddess sepulchre island friend berries
17	0.1446	Power of nature	pride earth woe throne form storm eye course thunders ray lightning thunder bones riflemen wreath misery storms realms wing bow

Figures 2 and 3 show the densities, which are output results of LDA for Topics 2 and 3, respectively^{2 3}. The density values indicate the proportion of each topic within a segment. Therefore, it could be suggested that the lower the proportion of the topic, the more prominent the presence of other topics in these segments. For the next discussion, segments were filtered out the top 65 segments to reduce the count from 1,041 segments in 593 poems. While most of the poems listed in Fig. 2 for Topic 2 are lyrical works, several poems are narrative poems, such as T33_THEOTT ('The Lady of Shalott', 1833), T62_ENODEN (*Enoch Arden*, 1862), and T69_COM-1K ('The Coming of Arthur', 1869). The number of narrative poems in Topic 3 is relatively lower. In addition, the proportion of Topic 3 in narrative poems, for instance T47_PRIESS (*The Princess*, 1947) and T42_AUDURT ('Audley Court', 1842), is also lower compared to other lyrical poems. The years of publication and writing of the poems in Topic 3 are scattered, but approximately half of the poems were published or written in the 1820s and 1830s.

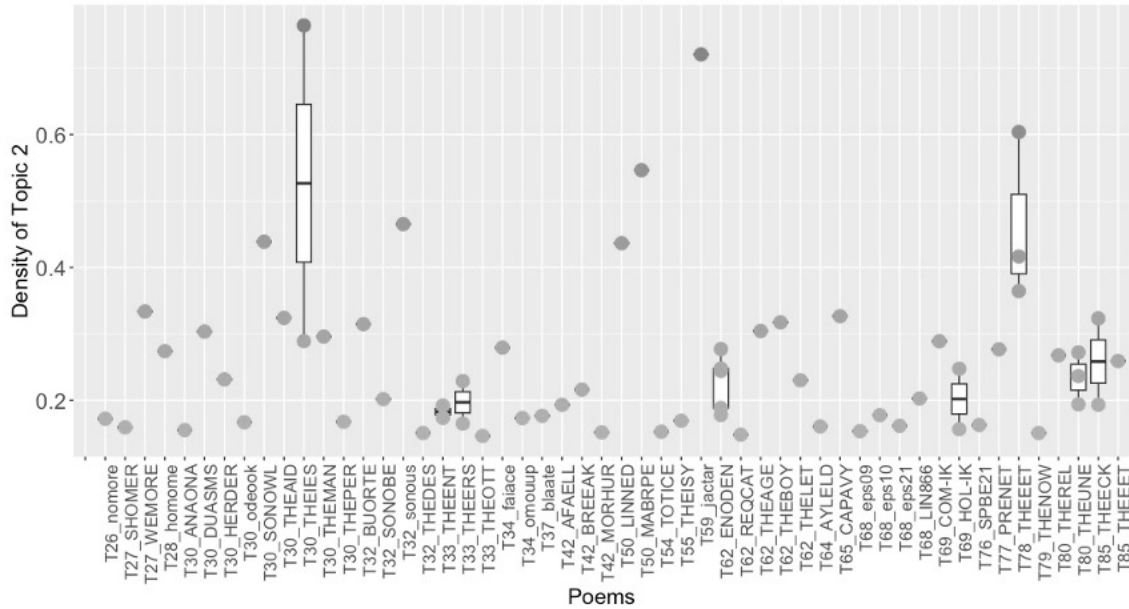


Fig. 2: Density plot of Topic 2.

Fig. 4 shows the mean density values of the top 65 segments for Topics 2 and 3. In cases where poems are duplicated in the top 65 segments, the mean density values were calculated to plot Fig. 4. It is worth noting that the density value for each poem is inversely proportional to the other. Specifically, as the density of Topic 2 increases, the density of Topic 3 decreases, or vice versa, in the poems for each topic. While most of the works tend to be allocated to only either Topic 2 or 3, there are some exceptions, such as T27_THEARP, T32_MINRIT, and T68_eps22.

² The abbreviated titles shown in lowercase represent that the poems are not published, whereas uppercase titles indicate the opposite. A 'T' at the beginning of the title indicates the poem is written by Alfred Tennyson, and the following two-digit numbers indicate the last two digits of the year of publication (or, if unpublished, of its composition).

³ For the comparison tables of the abbreviated and full titles of the poems under Topics 2 and 3, see the Appendices.

Based on the density values, it can be inferred that Tennysonian poems tend to depict sea and inland environments separately rather than concurrently and to the same extent in one work. The following section will elaborate on the content and details of Topics 2 and 3 and investigate the association between topics and the content of poems.

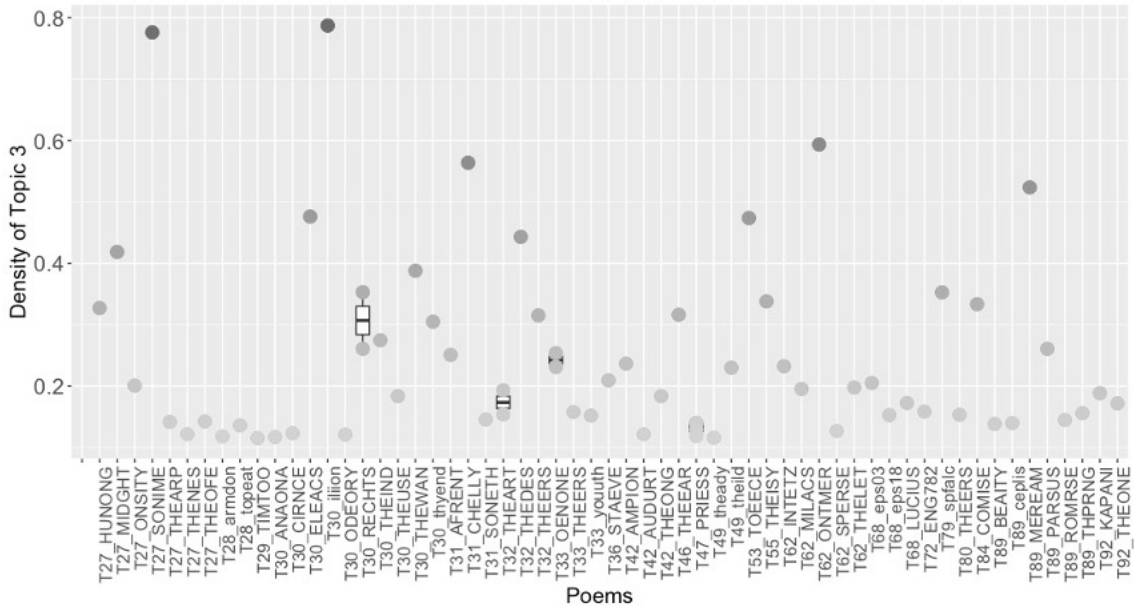


Fig. 3: Density plot of Topic 3.

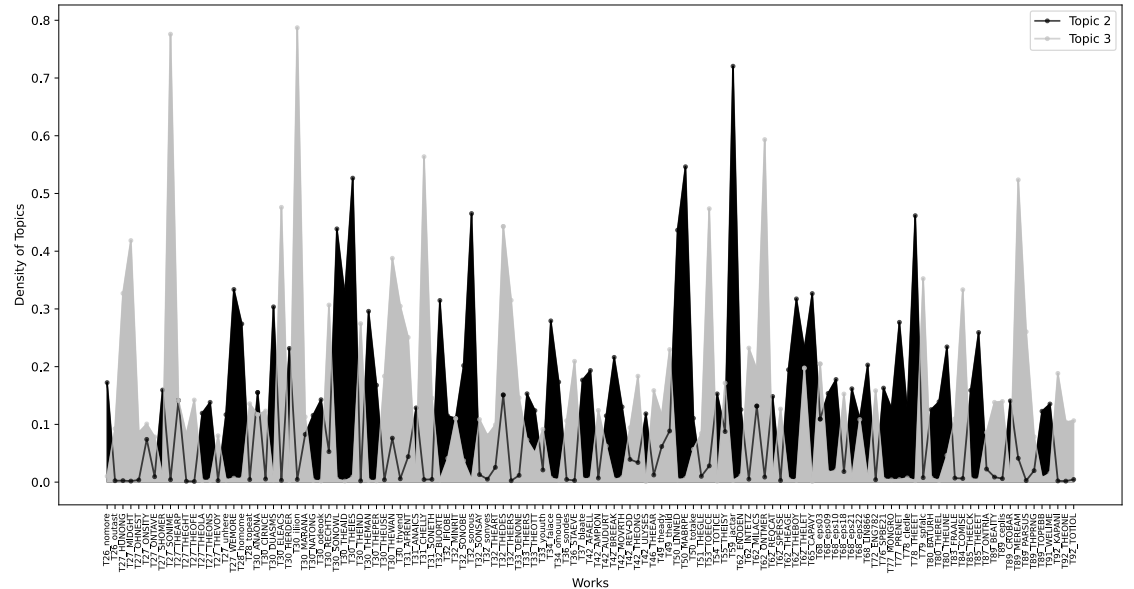


Fig. 4: Comparison of the mean value of density for Topics 2 and 3.

5. Discussion

Although the LDA analysis covers 18 topics that explain important and characteristic features of Tennyson’s poems, this paper and subsequent discussion focus solely on his descriptions of the natural environment in his poetry. As previous studies have pointed out, Tennyson had predilection for describing and exploring bodies of water in nature. On the subject of bodies of water, Topic 2 specialises in the sea and Topic 3 has keywords related to water, *river* and *lake* (see Fig. 5 for keywords of Topic 2 and Fig. 6 for Topic 3). Furthermore, detailed descriptions of bodies of water are depicted in various natural settings. In Topic 2, geographic features, such as *isle*, *shore*, *sand*, and *beach*, closely related to the sea, exquisitely represent landscapes. On the other hand, in Topic 3, plants and sounds, for instance, *pine*, *tree*, and *melody*, are ranked in the top 20 keywords. With these detailed elements, Tennyson succeeded in describing and conveying highly detailed scenes using words to illustrate the elements as well as the characters’ sentiments around the sea and river that appear in Topics 2 and 3.



Fig. 5: Word cloud of the top 20 keywords (Topic 2).

5.1. Sea

As Keirstead (2019: 75) notes, ‘Simply put, going to sea means almost certain death in Tennyson’. Death, dying people, and dying animals are mentioned directly, indirectly, and figuratively in 14 of Topic 2’s 65 prominent segments. Specifically, the word *sea* is associated with the concept of death in eight poems. However, in the assigned poems of Topic 2, the sea is not necessarily synonymous with death. The investigation of this topic reveals that the sea is linked to several concepts present in the assigned poems, including death. In ‘The Sailor Boy’ (1862; T62_THEBOY),

‘**danger** in the roaring **sea**’ (boldface added⁴) emphasises the dreadfulness of ‘a **devil**’ that ‘rises in my heart’, saying that this devil ‘in my heart’ is ‘[f]ar worse than any death to me’ (ll. 21–24). The final line of ‘Prefatory Sonnet’ (1877: l. 14; T77_PRENET) uses a metaphor to convey a sense of depth and gravity: ‘In **seas** of Death and sunless gulfs of Doubt’. Though the word *sea* is not used in ‘Sonnet [But were I Loved as I Desire to be]’ (1832; T32_SONOB), readers can easily connect the *surge* and *foam* to the sea in the following lines:

But were I loved, as I desire to be,
What is there in the great sphere of the earth,
And range of evil between death and birth,
That I should fear if I were loved by thee?

...

Apart upon a mountain, though the **surge**
Of some new deluge from a thousand hills
Flung **leagues** of roaring **foam** into the **gorge**
Below us, as far on as eye could see.

“Sonnet [But were I Loved as I Desire to be]” (1832: ll. 1–4, 11–14; T32_SONOB)
(Boldface and underlines added)

The final line of the abovementioned poem, ‘as far on as eye could see’ (l. 14) implies a gap between *us* and the surface of the seawater. The theme of distance between objects is a recurring motif in poems of Topic 2.

Related to distance and watercraft, both *boat* and *ship* also appear. These seagoing vehicles are associated with regions and areas that are beyond the sea and with the action of travelling to those far places. Because Britain is an island country, there can be no doubt that Tennyson associated the sea with going abroad or establishing relationships with other regions. Ships take people abroad as well as to sail into battle over the seas.

In ‘Lines [Here Often, When a Child, I Lay Reclined]’ (1850; T50_LINNED), a distant region (for Tennyson), Greece, is referred to in line 4 by the word *Grecian*, with the keyword of Topic 2, *ships*. Not only the description of Greece but also solitude, lamentation, and the distance of time are depicted in this poem.

In wartime in particular, the seas around the British Isles are the last line of defence for Great Britain. The most significant poem of Topic 2, the unpublished work ‘Jack Tar’ (1859 circa; T59_jactar) describes ‘the noble pride of the Mistress of the Seas’ (l. 4) in a battle, and the phrase ‘the **pride** of the **Mistress** of the Seas’ repeatedly appear in lines 4, 7, 11, 14, 18, and 21. ‘Jack Tar’ is considered to be written while ‘anticipating war with France’ (Ricks, 1987b: 604). In addition, the term ‘the Mistress of the Seas’ is the sobriquet of Great Britain; thus, it forms a further compelling connection between war/battle and the sea.

Battle is the most significant scene in ‘The Revenge: A Ballad of the Fleet’ (1878; T78_THEEET) and in ‘Buonoparte’ (1832; T32_BUORTE), which is regarding the Battle of Trafalgar. In these

⁴ Henceforth, boldface styling of terms in excerpts and quotations from poems will indicate the word to which the topic is assigned.

two poems, the sea is depicted as a site of battle. As with the images of shipwreck in *Enoch Arden* (1862; T62_ENODEN) and battles, the sea is associated with an inescapable, unwilling, and lonely death⁵. In scenes where the sea appears, the sound of its waves is almost the only sound to be heard. The openness of the sea and the absence of any natural or artificial surroundings are also signified. This nothingness around the characters in the scenes near the sea further amplifies the feelings of loneliness, despair and fear, especially when the characters in poems lose their loved ones.

Some works, including ‘Mablethorpe’ (1850; T50_MABRPE) and ‘To Professor Jebb’ (1889; T89_TOPEBB)⁶ describe foreign countries or regions without mentioning death. Not only are distant nations such as Egypt (‘To Professor Jebb’) conveyed in these poems but also their aloofness appears, specifically in the descriptions of archaic locations such as *Ilion*, recalling ancient Greece and Greek mythology.

Other poems under Topic 2 suggest mental, rather than physical, remoteness. ‘We Meet No More’ (1827; T27_WEMORE, quotation below) is a poem that both mental and physical remoteness depicted by referring to the parting of *I* who is standing ‘on a distant shore’, and *Ellen* (ll. 7, 12), and what separates them is ‘angry seas’ (l. 6). Although death is not explicitly represented in ‘We Meet No More’, the *[f]arewell* in the final line, in conjunction with the term ‘no more’, conveys a sense of mental remoteness, loneliness, and an enduring sense of eternity. The phrase ‘no more’ is used repeatedly in this poem, yet it does not directly refer to eternity or nothingness. Instead, it suggests that a particular event or state has reached its conclusion at a specific juncture and is not capable of reoccurrence in the future.

According to Huddleston and Pullum (2002: 831), the function of ‘no more’ and ‘no longer’ is that of absolute negators, relating to the concepts of stillness and to ‘any more/any longer’. Whilst *still* is regarded as a positively-oriented polarity-sensitive item (PPI), preferring positive contexts, ‘any more/any longer’ is characterised as a negatively-oriented polarity-sensitive item (NPI), preferring negative contexts. Huddleston and Pullum posit the view that both ‘any more’ and ‘any longer’ are aspectual NPIs. Despite the absence of overt expressions of time continuity in the poem, it does exhibit a consistent time flow that can be inferred from specific lines such as ‘[o]ur every hope on earth is past’ (l. 3), ‘[a]nd lonely thoughts of days gone o’er’ (l. 9), and ‘[f]arewell for ever, Ellen!’ (l. 12). The poem’s repetitive depiction of events serves to underscore the notion of time’s progression and its constancy. The absolute negators ‘no more’ and ‘no longer’ in the correspondence relationship of the aspectual NPI imply the negation of the continuity of time. Consequently, events or actions that transpired during the period of continuous time will cease to persist once the consistency of time is negated by the use of these words. It

⁵ The eponymous protagonist Enoch of *Enoch Arden* does not die in the shipwreck during his journey back home from the Far East but ends up dying in sorrow, deep love for his wife and family members, a profound sense of self-sacrifice and despair at the end of the story.

⁶ This poem is not ranked in the top 65 segments of Topic 2, but it is 86th and its density of Topic 2 terms is 0.1228.

can thus be interpreted that the phrase ‘we meet no more’, which appears in lines 5 and 11, is as ‘we will never see each other again’ or ‘we will never meet again’. The interpretations suggest that the protagonists had had opportunities to see each other up until a certain moment. However, it is implied that the occurrence of a future meeting is rendered implausible. This interpretation suggests a connotation of everlastingness in the phrase ‘no more’.

WE meet no more — the die is cast,
 The **chain** is broke that tied us,
 Our every hope on earth is past,
 And there’s no helm to guide us:
 We meet no more — the roaring **blast**
 And angry **seas** divide us!
 And I stand on a distant **shore**,
 The **breakers** round me swelling;
 And lonely thoughts of days gone o’er
 Have made this breast their dwelling:
 We meet no more — We meet no more:
Farewell for ever, Ellen!

‘We Meet No More’ (1827; T27_WEMORE)

The poem entitled ‘The First Quarrel’ (1880; T80_THEREL) also describes the separation of a man and a woman, and their fates are sealed when the man dies in the sea, as alluded to in the final line of the poem: ‘An’ the **boat** went down that night—the **boat** went down that night’ (l. 92). The narrative of ‘The First Quarrel’ (1880; T80_THEREL) is presented through the perspective of the *I*-narrator, who remains anonymous apart from being referred to as Harry’s wife. The narrative commences with the female narrator giving birth to a boy in challenging conditions. The narrator provides a detailed account of her circumstances to her physician, reflecting on her life in the context of her bereavement. Her husband, Harry, met his demise at sea. The narrator has been experiencing a state of unhappiness, precipitated by the demise of her spouse and her new born child. Nevertheless, the aspect that appears to be the most challenging for her is the fact that she lost her husband prior to having had the opportunity to reconcile with him following the very first and the last quarrel in their relationship. Their quarrel began when she suspected her husband of having an affair with another woman. Harry denied it, but the wife did not believe him. Before they become reconciled, Harry had to leave from the narrator’s presence for work. He had to travel to a location across the sea by boat. Following his departure, the wife realised that she was wrong, but it was too late, as the boat that Harry was on sank in the ocean, and she lost her husband and the opportunity to apologise to him at the same time. In ‘The First Quarrel’, therefore, the mental and physical distances, death, sea, loneliness and perpetuity, all of which are elements of Topic 2, are conveyed in a single poem. While ‘The First Quarrel’ serves to illustrate Keirstead’s (2019: 75) proposition, the poem also demonstrates that the sea in Tennyson is not exclusively associated with death; it is also linked to other elements, including mental and physical distances, loneliness, and eternity.

5.2. Mountain and River

Topic 2 associates the sea with distance, eternity, death, estrangement, foreign regions, and the past. In contrast, Topic 3 explores religious motifs and their impact on people's lives, as indicated by the keywords (Fig. 6). The top 20 keywords of Topic 3 associate with creeds as well as a focus on a moment or life before death but not everlasting or uncertain time and the afterlife. In terms of religious motifs, the most prominent keyword is *mountain* in Topic 3, which is more conspicuous than river. The religious mythologising of mountain-related terms in Tennyson, such as *pinnacles* in 'The Lotos-Eaters' (1833: l. 16) and *mountain* in 'Youth' (1833 circa), are referred to by Paden (1971: 157) and Ricks (1987a: 637).

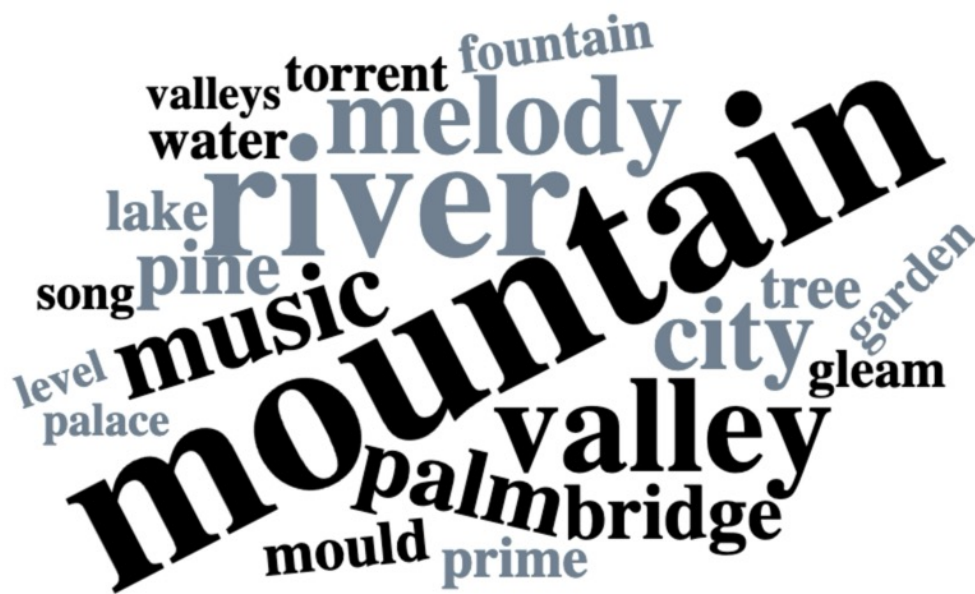


Fig. 6: Word cloud of the top 20 keywords (Topic 3).

In a poem, 'To E. L. on his Travels in Greece' (1853; T53_TOEECE), in which Topic 3 densely appears, the *Gods* and a *mountain* are not directly linked. However, a sense of holiness or, more likely, a sense of solemnity from the surrounding items is expressed through Topic 3's keywords.

Despite the absence of *god(s)* in 'Merlin and the Gleam' (1889; T89_MEREAM), the *mountain* are filled with *Wizard*, *Griffin*, *Giant*, *Fairies* and *dragons*, and other supernatural beings, giving the scene a surreal atmosphere. The definition of *Gleam*, a keyword in Topic 3 that frequently occurs in this work, is given by Tennyson himself as referring to 'poetic imagination' (Ricks, 1987c: 205). Neither *Gleam* nor 'poetic imagination' have any theological overtones or connections to the Almighty. However, *Gleam*, which appears on a mountain far from the real, human world, colours the otherworldly imaginary world of the work. In addition, the protagonist of this poem, Merlin 'follow[s] The **Gleam**' (l.10). A couple of stanzas later, 'The Master whisper'd' to Merlin, 'Follow The **Gleam**' (ll. 33–34). The presentiment of Merlin's death, as well as 'The Master', who suddenly appears in the poem and murmurs to Merlin, further adjoin the sacredness for the

word *Gleam*. In addition to their presence on the mountain, Tennyson depicted *gods* seated on the clouds (floating) on the heights. These deities and their elevated positions also bear connections to religious concepts. Another religious term, *Heaven*, also connotes altitude, for instance in ‘Thy Soul is Like a Landskip, Friend’ (1830; T30_thyend).

As the keywords in Fig. 6 indicate, *melody* and sound are widely spread in the poems of Topic 3. Within the most dominant works of Topic 3, ‘Ilion, Ilion’ (1830; T30_iliion), ‘Elegiacs’ (1830; T30_ELEACS), ‘The Dying Swan’ (1830; T30_THEWAN) and others, are filled with melodies, sounds, and songs. In the 19 lines of ‘Ilion, Ilion,’ the word *melody* appears five times: ‘when wilt thou be **melody** born?’ (ll. 2, 9, 12, 19) and ‘ever onward to a **melody**’ (l. 5). Moreover, the first line of ‘Elegiacs’, ‘Lowflowing breezes are roaming’ leads to a passage in which living creatures make sounds: ‘Barketh the **shepherd-dog** cheerily; the grasshopper carolleth clearly; / Deeply the **turtle** coos; shrilly the **owlet** halloos’ (ll. 5–6). One theme of ‘The Dying Swan’ is undoubtedly death, but this poem is unlike the poems of Topic 2 with their scattering songs and sounds:

With an inner voice the **river** ran,
 Adown it floated a dying **swan**,
 And loudly did lament.
 ...
 One willow over the **water** wept,
 And shook the **wave** as the wind did sigh;
 ...
 The wild **swan**’s **death-hymn** took the soul
 Of that waste place with joy
 Hidden in sorrow: at first to the ear
 The **warble** was low, and full and clear;
 ...
 With a music strange and manifold,
 Flow’d forth on a carol free and bold;
 As when a mighty people rejoice
 With **shawms**, and with **cymbals**, and harps of gold,
 And the tumult of their **acclaim** is roll’d
 ...
 And the creeping mosses and clambering weeds,
 And the willow-branches hoar and dank,
 And the wavy **swell** of the souging reeds,
 And the wave-worn horns of the echoing **bank**,
 And the silvery **marish-flowers** that throng
 The desolate **creeks** and **pools** among,
 Were flooded over with eddying **song**.
 ‘The Dying Swan’ (1830: ll. 5–7, 14–15, 21–24, 29–33, 36–42; T30_THEWAN)

As can be seen, the anthropomorphic plants and animals talk, sing, and surround the people in the poems of Topic 3. Unlike the cases of Topic 2, living creatures are present in the poems and make various sounds. These works do not convey any impression of fear or despair of the type

felt in the works of Topic 2. Furthermore, both in 'The Dying Swan' and in the aforementioned 'Merlin and the Gleam', the final moments of the characters' deaths are not described. These dying animate beings are surrounded by sounds and other living beings. These situations do not necessarily convey utter solitariness or despair to readers. This therefore suggests not only the reflection of the characters' emotions but also Tennyson's intention to sketch differently in relation to the natural surroundings, evoking these emotions not directly but in an indirect way.

Not only incidental sounds but also those of people, animals, and apostrophised nature, including plants and water, crowd the poems of Topic 3; this contrasts with Topic 2 poems whose scenes highlight the sea. Even the difference in altitude alluded to by mountains and valleys influences how people perceive their surroundings. In some cases, the objects under discussion are located distant in the physical realm of poems; however, in the case of Topic 3, these objects are rendered visible through their depiction in poems. This visibility, whether clear or vague, can be seen by the narrators, the characters, and the readers, even if they are distant. Such visibility does not convey the sense of despair felt in the poems of Topic 2. This suggests that the objects not only mirror the characters' emotions but also that Tennyson's intention was to sketch and evoke these emotions indirectly rather than directly.

In Topic 3's poems, Tennyson employs rivers as a metaphor for life, for example in 'Compromise' (1884; T84_COMISE):

Steersman, be not precipitate in thine act
Of **steering**, for the **river** here, my friend,
Parts in two channels, moving to one end—
This goes straight forward to the **cataract**:
That streams about the **bend**;
But though the **cataract** seem the nearer way,
Whate'er the crowd on either **bank** may say,
Take thou the '**bend**,' 'twill save thee many a day.

'Compromise' (1884; T84_COMISE)

As can be seen in the quotation of 'Compromise' above, rivers wind through bends and cataracts until they finally reach a broad, soundless, open sea. This suggests that lives with ups and downs ultimately come to an end. Poems in Topic 3 do not extend rivers to the ocean, focusing instead on the present moment rather than the eternity or death of the sea, through bends. Religious connotations pertaining to death, such as the term *Heaven*, are not necessarily taken as an indication of a character's death within a poem or of characters' eternal life after death. Rather, the concept associates with a belief that one may hold while still alive. Hence, the poems in Topic 3 centre on the moments when the characters are alive. The belief and the focus on the living moment serve as the pivot in the poems of Topic 3.

5.3. Did LDA overlook a canonical poem?

Discussion of Tennyson's concept of the river cannot overlook several works, for instance, 'The Lady of Shalott'. The scene in which the Lady is on a boat floating into Camelot is illustrative. The river in this poem is one of the impressive elements in many paintings of the scene, such as those by J. W. Waterhouse and Henry Peach Robinson (Cheshire, 2009: 57, 90–91). However, this poem does not appear in the discussion of the river in this section or in the top 65 significant segments of Topic 3. Did LDA miss a prominent and representative theme in the Tennysonian canon, while it was handling and analysing large amounts of poetry, including the 'great unread' dimensions?

'The Lady of Shalott', belongs in fact to Topic 7 (the eighth most significant poem/segment) rather than to Topic 3. The reason why this poem was assigned to the topic is that it emphasises the extent of the Lady as well as the Lady's despair and her remembrance, even following her death. Here again, the river metaphorically suggests the Lady's life. In the poem, the river does not reach a vast, silent sea after all of its bends, windings and cataracts through Camelot; nevertheless, rivers generally reach the sea finally as vicissitude-filled lives inevitably come to an end. The river in 'The Lady of Shalott' unquestionably attracts readers' attention; however, what LDA detected here was not the Lady's life rendered as the river but her melancholy, forlornness and despair.

Tennyson, as a 'landscape-painter in words, a colourist' (Hair, 1991), lays out his multiple themes in detailed word portraits of nature. Although some LDA results contradict earlier scholarly readings and readers' intuitions, as shown in this section, LDA unfolded and spotlighted latent semantic connections between words spread over multiple poems that are hidden behind canonical or eye-catching themes or motifs. However, this does not amount to a cancellation by LDA of the outcomes of the close readings found in previous studies.

6. Conclusion

In this study, I observed poetic depictions of sea and inland landscapes, mountains and rivers based on the results of the LDA topic model. What were mentioned and discovered in this study were not necessarily new and groundbreaking findings; nevertheless, they revealed that quantitative and qualitative approaches combined can support each other and help to discover hitherto unknown dimensions of Tennyson. Of course, I was not insisting that poems should have only one topic. The topics in LDA are interrelated; in particular, in the works with two topics, they are often complexly intertwined rather than occurring independently within each work; this is not a problem but rather a testimonial to how complex Tennyson's depictions are in his works and how accurately the topic model captures this complexity.

In addition to the interpretations of the sea in Section 5.1, eternity can interweave with immortality, which is another prominent theme in Tennyson. Although LDA facilitated my reading of Tennyson's poems beyond his canon and revealed new dimensions, this paper does not provide a comprehensive analysis of all their elements. Future research should explore Tennyson's works

in greater depth, using both qualitative and quantitative approaches to investigate the motifs of immortality and others that emerged from the LDA results. By doing so, it will be possible to prevent leaving a gap of the ‘great unread’ in Tennyson’s works.

Acknowledgements

This paper is partly based on the author’s presentation at PALA 2022 *Style and Sense(s)*, July 8, at Aix-Marseille Université, and the 118th conference of Japan Society of Stylistics, June 18, at Kyorin University. I would like to thank the chairs and audience members for useful discussion.

This paper is in part based on the author’s doctoral dissertation submitted to Osaka University in November 2023.

The research was in part supported by the Japan Science and Technology Agency Support for Pioneering Research Initiated by Next Generation [grant number JPMJSP2138], to which the author of the present article was granted funding from October 2021 to March 2023.

Bibliography

- [1] Blei, M.D., Ng, Y.A. and Jordan, I.M. (2003) Latent dirichlet allocation. *Journal of Machine Learning Research*, 3: 993–1022.
- [2] Brooke, A.S. (1894) *Tennyson: His Art and Relation to Modern Life*. London: Isbister.
- [3] Chesire, J. (Eds.) (2009) *Tennyson Transformed: Alfred Lord Tennyson and Visual Culture*. Burlington: Lund Humphries.
- [4] Dixon, M.W. (1896) *A Primer of Tennyson*. London: Methuen & Co.
- [5] Fujita, I. (2022) On segment size in poetry analysis: using the latent dirichlet allocation method. *The Japanese Journal of Digital Humanities*, 3(1): 3–15.
- [6] Fujita, I. (2023) ‘To spread the Word by which himself had thriven’: Analysis of Alfred Tennyson’s use of language based on the LDA topic model. *English Corpus Studies*, 30: 3–26.
- [7] Fulweiler, W.H. (1965) Tennyson and the “summons from the sea”. *Victorian Poetry*, 3(1): 25–44.
- [8] Hair, S.D. (1991) *Tennyson’s Language*. Toronto: University of Toronto Press.
- [9] Henrichs, A. (2019) Deforming Shakespeare’s sonnets: topic models as poems author(s). *Criticism*, 61(3): 387–412.
- [10] Huang, C. (2020a) Chugoku-no misuterii shousetsuniokeru topikku kaiseki-no kokoromi. (Experimental topic analysis on Chinese mystery prose texts). *Studies in Language and Culture*, 29 : 1–17.
- [11] Huang, C. (2020b) Chugoku-no misterii shousetsu-wo meguru keiryouteki bunseki: Tei Shou Sei to Ki Ba Sei-no sakuhin-wo chuushinn-ni. (Quantitative analysis of Chinese mystery novels: focusing on the works of Cheng Xiao Qing and Gui Ma Xing). *Text mining and Digital Humanities*, 2020: 31–45.
- [12] Huddleston, R., & Pullum, G. K. (2002) *The Cambridge grammar of the English language*. Cambridge University Press.
- [13] Imao, Y. (2022) CasualConc, version 2.0.8. Available (newer version) at: <https://sites.google.com/site/casualconcj/casualconc/CasualConc>.
- [14] Iwata, T. (2015) *Topic Model*. Tokyo: Kodan-sha.
- [15] Jockers, L.M. (2014) *Text Analysis with R for Students of Literature*. Heidelberg, New York, Berlin, London: Springer Cham.
- [16] Keirstead, C.M. (2019) Sea dreams and realities: Tennyson, travel, and the shifting currents of littoral space in Victorian culture. *Victorian Poetry*, 57(1): 73–99.
- [17] Kiyama, N. (2018) How have political interests of U.S. Presidents changed?: a diachronic investigation of the State of the Union addresses through topic modeling. *English Corpus Studies*, 25: 79–99.
- [18] Kuroda, A. (2017) Quantitative analysis of literary works: novels of Sir Arthur Conan Doyle. *Text mining and Digital Humanities*, 2017: 23–41.

- [19] Matsukawa, H., Oyama, M., Negishi, C., and Arai, Y. (2018) Analysis of the free descriptions obtained through course evaluation questionnaires using topic modeling. *Japan Journal of Educational Technology*, 41(3): 233–244.
- [20] McCallum, A.K. (2002) Mallet: a machine learning for language toolkit. Available at: <http://mallet.cs.umass.edu>.
- [21] Moretti, F. (2013) *Distant Reading*. London: Verso.
- [22] Navarro-Colorado, B. (2018) On poetic topic modeling: extracting themes and motifs from a corpus of Spanish poetry. *Frontiers in Digital Humanities*, 5: 5–15.
- [23] Okabe, M. (2019) *Thou and You* in Emily Dickinson’s poems using topic modeling: reconsideration of interjections. *Proceedings of the Japanese Association for Digital Humanities Conference 2019*: 125–131.
- [24] Okazawa, T. (1969) A study of Tennyson’s predilection for waters. *Bulletin of Chofu Gakuen Women’s Junior College*, 2: 1–12.
- [25] Onodera, D., Huang, L. and Yoshioka, M. (2016) Classification of new article by using facet-biased topic model and distance metric learning. *The 30th Annual Conference of the Japanese Society for Artificial Intelligence 2016*: 1–4.
- [26] *Oxford English Dictionary*. Available at: <https://www.oed.com> (Last accessed 22 October 2023).
- [27] Paden, W.D. (1942) *Tennyson in Egypt: A Study of the Imagery in his Earlier Work*. University of Kansas.
- [28] Plamondon, M.R. (2005) Computer-assisted phonetic analysis of English poetry: a preliminary case study of Browning and Tennyson. *The Journal of Computer Text Processing*, 14(2): 153–175.
- [29] Poets Series (2013) *Alfred, Lord Tennyson*. East Sussex: Delphi.
- [30] Rhody, M.L. (2012) Topic modeling and figurative language. *Journal of Digital Humanities*, 2(1): 19–35.
- [31] Ricks, C. (1969) *The Poems of Tennyson*. London: Longman.
- [32] Ricks, C. (1987a) *The Poems of Tennyson*. 2nd ed, vol. I. London: Longman.
- [33] Ricks, C. (1987b) *The Poems of Tennyson*. 2nd ed, vol. II. London: Longman.
- [34] Ricks, C. (1987c) *The Poems of Tennyson*. 2nd ed, vol. III. London: Longman.
- [35] Shaw, W.D. (1973) *Tennyson’s Style*. Ithaca and London: Cornell University Press.
- [36] Tabata, T. (2017) Mapping Dickens’s novels in a network of words, topics, and texts: topic modelling a corpus of classic fiction. *Japanese Association for Digital Humanities Conference 2017*, Doshisha University.
- [37] Tabata, T. (2018) Mapping Dickens’s novels in a network of words, topics, and texts. *Text Mining and Digital Humanities, 2017*: 51–60.
- [38] Tabata, T. (2020) Latent topics in British classic fiction: using LDA to classify texts into meaningful groups. *Text Mining and Digital Humanities, 2019*: 47–58.
- [39] Tennyson, C. (1949) *Alfred Tennyson*. London: Macmillan.
- [40] Tennyson, H. (1897) *Alfred Lord Tennyson: A Memoir by His Son*. vol. I–III. London: Macmillan.
- [41] Thomas, J. (2019) *Tennyson Echoing Wordsworth*. Edinburgh: Edinburgh University Press.

Appendices

Appendix I: Comparison table of abbreviated and full-length titles of poems: Topic 2.

Abbreviated Titles	Year of Publication	Poem Titles
T26_nomore	1826	'No More'
T27_SHOMER	1827	'Short Eulogium on Homer'
T27_WEMORE	1827	'We Meet No More'
T28_homome	1828	'Home'
T30_ANAONA	1830	'Anacaona'
T30_DUASMS	1830	'Dualisms'
T30_HERDER	1830	'Hero to Leander'
T30_odeook	1830	'Ode: O Bosky Brook'
T30_SNOWL	1830	'Song: The Owl'
T30_THEAID	1830	'The Mermaid'
T30_THEIES	1830	'The Sea Fairies'
T30_THEMAN	1830	'The Merman'
T30_THEPER	1830	'The Grasshopper'
T32_BUORTE	1832	'Buonoparte'
T32_SONOB	1832	'Sonnet [But were I Loved As I Desire to be]'
T32_sonous	1832	'Sonnet [Conrad! Why Call Thy Life Monotonous?]'
T32_THEDES	1832	'The Hesperides'
T33_THEENT	1833	'The Lovers Tale: A Fragment'
T33_THEERS	1833	'The Lotos Eaters'
T33_THEOTT	1833	'The Lady of Shalott'
T34_faiace	1834	'Fair is that Cottage its Place'
T34_omouup	1834	'O Mother Britain Lift thou Up'
T37_blaate	1837	'Black Bull of Aldgate'
T42_AFAELL	1842	'A Farewell'
T42_BREEAK	1842	'Break, Break, Break'
T42_MORHUR	1842	'Morte D'Arthur'
T50_LINNED	1850	'Lines [Here Often, When a Child, I Lay Reclined]'
T50_MABRPE	1850	'Mablethorpe'
T54_TOTICE	1854	'To the Rev. F.D. Maurice'
T55_THEISY	1855	'The Daisy'
T59_jactar	1859	'Jack Tar'
T62_ENODEN	1862	Enoch Arden
T62_REQCAT	1862	'Requiescat'
T62_THEAGE	1862	'The Voyage'
T62_THEBOY	1862	'The Sailor Boy'
T62_THELET	1862	'The Islet'
T64_AYLELD	1864	'Aylmer's Field'
T65_CAPAVY	1865	'The Captain (A Legend of the Navy)'
T68_eps09	1868	'Epigrams' (1868, ix)
T68_eps10	1868	'Epigrams' (1868, x)
T68_eps21	1868	'Epigrams' (1868, xxi)
T68_LIN866	1868	'Lines (1865–1866)'
T69_COM-IK	1869	'The Coming of Arthur'
T69_HOL-IK	1869	'The Holy Grail'
T76_SPBE21	1876	'Songs from the Plays ['Becket' Act II Scene I]'
T77_PRENET	1877	'Prefatory Sonnet'
T78_THEEET	1878	'The Revenge: A Ballad of the Fleet'
T79_THENOW	1879	'The Defence of Lucknow'
T80_THEREEL	1880	'The First Quarrel'
T80_THEUNE	1880	'The Voyage of Maeldune'
T85_THEECK	1885	The Wreck'
T85_THEEET	1885	'The Fleet'

Appendix II: Comparison table of abbreviated and full-length titles of poems: Topic 3.

Abbreviated Titles	Year of Publication	Poem Titles
T27_HUNONG	1827	‘Huntsmans Song’
T27_MIDGHT	1827	‘Midnight’
T27_ONSITY	1827	‘On Sublimity’
T27_SONIME	1827	‘Song it is the Solemn Even Time’
T27_THEARP	1827	‘The Exiles Harp’
T27_THENES	1827	‘The Vale of Bones’
T27_THEOFE	1827	‘The Dell of E—’
T28_armdon	1828	‘Armageddon’
T28_topeat	1828	‘To Poesy [O God, Make This Age Great]’
T29_TIMTOO	1829	‘Timbuctoo’
T30_ANAONA	1830	‘Anacaona’
T30_CIRNCE	1830	‘Circumstance’
T30_ELEACS	1830	‘Elegiacs’
T30_iliion	1830	‘Ilion, Ilion’
T30_ODEORY	1830	‘Ode to Memory’
T30_RECHTS	1830	‘Recollections of the Arabian Nights’
T30_THEIND	1830	‘The Poets Mind’
T30_THEUSE	1830	‘The Deserted House’
T30_THEWAN	1830	‘The Dying Swan’
T30_thyend	1830	‘Thy Soul is Like a Landskip, Friend’
T31_AFRENT	1831	‘A Fragment’
T31_CHELLY	1831	‘Check Every Outflash Every Ruder Sally’
T31_SONETH	1831	‘Sonnet [Me My Own Fate to Lasting Sorrow Doometh]’
T32_THEART	1832	‘The Palace Oo Art’
T32_THEDES	1832	‘The Hesperides’
T32_THEERS	1832	‘The Sisters’
T33_OENONE	1833	‘(E none’
T33_THEERS	1833	‘The Lotos Eaters’
T33_youuth	1833	‘Youth’
T36_STAEVE	1836	‘St Agnes’ Eve’
T42_AMPION	1842	‘Amphion’
T42_AUDURT	1842	‘Audley Court’
T42_THEONG	1842	‘The Poet’s Song’
T46_THEEAR	1846	‘The Golden Year’
T47_PRIESS	1847	The Princess
T49_theady	1849	‘The Little Lady’
T49_theild	1849	‘The Losing of the Child’
T53_TOEECE	1853	‘To E.L. on His Travels in Greece’
T55_THEISY	1855	‘The Daisy’
T62_INTETZ	1862	‘In The Valley of Cauteretz’
T62_MILACS	1862	‘Milton (Alcaics)’
T62_ONTMER	1862	‘On Translations of Homer’
T62_SPERSE	1862	‘Specimen of a Translation of the Iliad in Blank Verse’
T62_THELET	1862	‘The Islet’
T68_eps03	1868	‘Epigrams’ (1868, iii)
T68_eps18	1868	‘Epigrams’ (1868, xviii)
T68_LUCIUS	1868	‘Lucretius’
T72_ENG782	1872	‘England and America in 1782’
T79_spfalc	1879	‘Songs From The Plays [‘The Falcon’]’
T80_THEERS	1880	‘The Sisters’
T84_COMISE	1884	‘Compromise’
T89_BEAITY	1889	‘Beautiful City’
T89_ceplis	1889	‘Cephalis’
T89_MEREAM	1889	‘Merlin and the Gleam’
T89_PARSUS	1889	‘Parnassus’
T89_ROMRSE	1889	‘Romneys Remorse’
T89_THPRNG	1889	‘The Progress of Spring’
T92_KAPANI	1892	‘Kapiolani’
T92_THEONE	1892	‘The Death of Ænone’