

「論文」

**“Dictions” by *Two Brothers*,
Charles and Alfred Tennyson**

Iku FUJITA

Abstract

This study employs a quantitative method, latent Dirichlet allocation topic model, to examine the distinctive thematic and lexical characteristics of poems by Alfred, Lord Tennyson, and his brother, Charles Tennyson, focusing primarily on *Poems by Two Brothers* (1827), the first published collection by Alfred, Lord Tennyson. It has been said that there is some ambiguity surrounding the poems’ authorship within the collection, originally published without annotations indicating each poem’s author. This research leverages topic model to uncover patterns in the diction and thematic inclinations of the two brothers. The results of the latent Dirichlet allocation analysis indicate that each poet gravitated toward certain specific topics as dominant themes in his works. Topic 17, which is associated with romantic sentiments and primarily physical descriptions of women, is prevalent in the poems written by Charles Tennyson, while Topic 1, which concerns themes of masculinity, enthusiasm, and battle, is prominent in the poems written by Alfred Tennyson. These findings highlight the distinctive differences between the two brothers in word usage. It is noteworthy that this study represents a novel application of topic model in examining characteristic diction of the two poets, offering internal evidence of the distinct word usages within the Tennysons’ collaborative collection. This exploration underscores that topic model is effective in distinguishing thematic tendencies as well as the characteristic diction of the two authors, Alfred and Charles Tennyson.

1. Introduction

This study investigates the characteristic content of two poets, Alfred, Lord Tennyson (1809–1892) and his brother, Charles Tennyson Turner (1808–1879), using a

quantitative method, the latent Dirichlet allocation (LDA) topic model (Blei et al., 2003). *Poems by Two Brothers* (1827)¹ is the first collection of 19th-century poet, Alfred Tennyson and his brother, Charles Tennyson. The authors of the poems in the collection had been considered two, Alfred and his brother, Charles, when *Poems by Two Brothers* was published. There was, however, no signature or annotation that referred to the author of each poem. The second edition of the collection was published in 1893, soon after the death of Alfred. The most significant change from the first to the second edition was the capital letters at the end of each poem suggesting its author. These notations clarify not only which poems were written by Charles or Alfred but also suggest the cooperation of another author, Frederick; interestingly, some works remain unidentified.

The authorship attribution of the *Poems by Two Brothers* as well as the collection itself have attracted the attention of few scholars. Brimley (1972) mentions that Alfred is responsible for roughly half of the poems in the collection. Paden (1964) qualitatively challenges the matter of unidentified authors by comparing capital letter annotations and notes written in the two manuscripts (Haddelsey's and Charles's copy) and the second edition of *Poems by Two Brothers*. Paden (1964) further identifies the authorship of 18 poems out of 20 poems, whose authorship are marked as uncertain or doubtful poems and for which the initials on two manuscripts and the 1893 edition do not match. Although there still remains two poems with uncertain authorship, it seems there is no study to follow Paden, perhaps because he had this to say about the collection as a whole: "None of the poems in question has any noticeable literary value, to be sure" (Paden, 1964, p. 147). This low estimation of the poems dovetails with criticism that the collection is largely imitations of fashionable styles at the time (Delphi Poets Series, *Alfred, Lord Tennyson*, 2013). In fact, successive studies on Alfred, Lord Tennyson do not fully address the authorship of poems in the collection, and it can be assumed that their perceived lower literary value might have lessened the interest of other scholars. Several collections of Alfred Tennyson's have been published, many of them edited by Christopher Ricks. The second edition of *The Poems of Tennyson* published in 1987 is compiled in the three volumes. Ricks lists six poems as "doubtful poems: poems attributed to Tennyson [Alfred]" in Appendix C of the third volume. Five of six poems, "Egypt," "The Deity," "On the Moon-Light Shining upon a Friend's Grave," "The Dying Christian," and "Switzerland," are marked as doubtful or uncertain

works in the 1893 edition of the *Poems by Two Brothers* (Ricks, 1987, pp. 641–646). In the 1893 edition, the sixth poem, “Song [To Sit Beside a Christal Spring],” was assigned “A.T.,” suggesting it is Alfred’s work. This conclusion differs from that of Paden (1964), who claimed that the five poems except for “Egypt” are attributed to Charles.

Ricks states that authorship attribution should be more concerned with external evidence, such as annotation in manuscripts than with internal evidence, such as poetic diction (1987, p. 647). Nonetheless, depending solely on the external evidence is limited by the absence of reliability and/or of information in the collection’s manuscripts. Paden (1964) considered the handwriting and annotations in the manuscripts, as well as the differences in style but found few explanations for specific stylistic elements that could be used to identify the authors of particular works.

Regarding stylistics, function words such as determiners, prepositions, and conjunctions are frequently highlighted. It is notable that quantitative authorship attributions as well as stylometry (quantitative stylistic analysis), often prioritize the examination of function words, or the most frequent words in a text, in their analysis of prose works. In contrast, content words (e.g., nouns, verbs, adjectives, and adverbs) receive comparatively less attention in authorship attribution studies. The rationale behind this is that the choice of content words is considered to reflect the content of the works, rather than the author’s style. An analysis of poetic style will consider rhyme, meter, and rhythm. In particular, in rhyming (and other sound effects), syllable rhyme is constituted less by function words than by content words. Given this, it can be posited that the word usage and figurative expressions using content words, namely poetic diction, are related to the characteristics of poems by Alfred and Charles.

Several previous studies using quantitative approaches on the works of Alfred Tennyson’s reveal that the 1820s poems use adverbs and nouns differently from other works written in the 1830s or later (Fujita, 2020, 2023). Fujita (2020) suggests a chronological difference in Alfred’s use of *-ly* adverbs using Correspondence Analysis. Fujita (2023) utilizes LDA topic model to analyze Alfred’s use of nouns. Fujita found that certain topics are concentrated in the 1820s poems. Both of Fujita’s studies indicate that the authorship differences possibly caused the results, although Fujita (2023) did not separate Charles’s and Alfred’s works in her analysis.

The analysis of content words and/or function words, namely poetic diction, can

be regarded as an examination of internal factors. This perspective opposes Ricks' suggestion. Nevertheless, if the analysis of this study can demonstrate discrepancies in internal elements through quantitative analyses, the results can help to augment the evidence found in previous studies and elucidate the divergences among authors. The use of a quantitative method engenders objective aspects and provides a divergent perspective from that of internal and external evidence. This study therefore employs a quantitative approach, the LDA topic model (Blei et al., 2003), which is adept in detecting semantic structures in the text data. This study aims to address two research questions: 1) Can LDA detect the differences in poetic diction between the works of Alfred and Charles? And 2) If LDA detects differences, what characteristics do the two authors exhibit? To investigate these questions, the study considers content words, which fill semantic roles and are intimately associated with poetic diction.

2. Data and methodology

2.1 Data

The target dataset (corpus) comprises 525 poems. The statistical description of the corpus is shown in Table 1. Among the 525 poems, 102 are from the first edition of *Poems by Two Brothers*, and 24 are unpublished works but are assumed to have been written in the 1820s by Alfred. The remaining 399 poems are Alfred's lyrical poems published/written in the 1830s to the 1890s. Each poem was assigned to a single file with filenames to indicate the author names, publication years, and abbreviated poem titles. The authorship of the poems, number of works for each category, and filename examples are shown in Table 2. The authorship of the poems from *Poems by Two Brothers* refers to the annotations in the second edition of *Poems by Two Brothers* (1893). The question mark ("?",) with the initial of author's names (e.g., "A.T. (?)") suggests doubtful authorships. The initials of both Alfred and Charles (e.g., "A.T. or C.T.") indicates that it is doubtful which of Alfred or Charles is the author of the poems ("The Deity" and "The Dying Christian"). The 1893 edition proposed the existence of another author, Frederick, which the "A.T. or C.T." annotation further indicates was not the case, as the poems were not written by Frederick. "The Egypt" is the only poem with the note, "Begun C.T., finished A.T.," seemingly indicating that the work was begun and mostly written by Charles, with Alfred writing the last two stanzas (Paden,

1964).

The target corpus included the works from other collections and publication dates other than 1827. Analyzing poems not included in *Poems by Two Brothers* made it possible to ascertain whether any characteristics observed in Alfred’s works were limited to a specific date or poems or if they were more pervasive throughout his career. While Alfred has a prolific oeuvre of over 50 narrative poems, this study excludes such works from the analytical purview. This is due to the fact that Tennyson’s *Poems by Two Brothers* (1827) is composed exclusively of lyrical poems. The notable distinctions between Alfred’s lyrical and narrative poems extend beyond the length of each work. They also encompass the themes, content, and characters present in the poems. It is therefore to be understood that narrative poems, while presented in the format of poetry, will exhibit content that is more closely aligned with that of prose. This discrepancy in word usage was noted by Fujita (2023). The subsequent analysis

Table 1. Descriptive statistics of the target corpus

Number of poems	525
The total tokens	168,282
Minimum number of words per poem	12
Maximum number of words per poem	18,662
Average number of words	320.54

Table 2. Breakdown of poems in the corpus and their filenames

	Authors* ¹	Number of works	Filenames
Poems in <i>Poems by Two brothers</i> (1827)	Alfred	46	A27_TITLE* ²
	Alfred (?; doubtful)	4	a27_TITLE
	Charles	48	C27_TITLE
	Charles (?; doubtful)	1	c27_TITLE
	Alfred or Charles Charles started and Alfred finished (mostly Charles)	2	ac27_TITLE
Poems in <i>Poems by Two brothers</i> (1893) but not in the 1827 edition	Alfred	3	Aup_title* ³
Unpublished poems but thought to have been written in 1820s	Alfred	21	T20_title* ^{3*4}
Lyrical poems of Alfred Tennyson published in the 1830s–1890s	Alfred	399	e.g., T30_TITLE/title* ^{3*4}

*1 For the works from *Poems by Two Brothers*, the authors assigned in the 1893 edition are listed.

*2 Abbreviated titles of poems are inserted where TITLE suggests.

*3 The unpublished poems’ abbreviated titles are indicated by lower case letters.

*4 The two digits following T indicate the last two digits of the publication year.

will therefore focus on Alfred Tennyson's lyrical poems to minimize the potential impact of genre on the results.

2.2 Methodology

The use of LDA allowed researchers to detect possible semantic links within the corpus under study and to categorize words that appeared in numerous documents into topics (Tabata, 2018, p. 52). The term *topic* here refers to a group of words, but it is not the same as the meaning used in field such as linguistics and literature. This method is considered to be well suited for content words, such as nouns that appear in a work's content and verbs that describe the movements of characters, because it uncovers latent semantic links among words based on their tendencies to co-occur. LDA employs the concept of the "bag of words," an approach that considers each document in the corpus as a bag (Jockers, 2014, p. 137). The larger the bag, the greater the likelihood of discovering words that are likely to co-occur within the same bag. The variance in document size is equal to that of bag size. With regard to the analysis of prose texts, Jockers posits that it is beneficial to divide novels (and other voluminous documents) into segments and then run the model (2014, p. 137). He asserts that LDA captures both overarching themes that traverse the entirety of a novel and themes that emerge and then recede at particular points.

Regarding segment size, namely the number of terms that compose a bag, Fujita (2022) proposed a relevant segment size for LDA practice in poetry analysis. This paper makes reference to Fujita (2022) to employ the mean token value (320 words, as illustrated in Table 1) of the corpus in question. The tokens in each poem are enumerated in sequence from the initial token to the concluding token, thereby segmenting each document. Once each text was divided into equal-sized consecutive segments, the two final parts were joined together, unless the final segment exceeded 160 words in length (which is the same as the half of the segment size). In the case of poems with a total number of tokens below 320, no duplicate segments were created; instead, each poem was treated as a single segment. Thus, the largest segment size was 480 words, while the smallest was 12 words. The 525 poems in the target corpus were split into 995 segments, which were subjected to further analysis.

All words in the texts were assigned part-of-speech tags employing a tag set CLAWS5, as given in the British National Corpus (Leech & Smith, 2000). Because

content words are more likely than function words to capture the ideational content of a text and poetic diction as well as to convey the poets’ mindsets, the current study limited its focus to nouns, verbs, and adjectives. Adverbs were excluded from the analysis because some of them do not convey semantic elements; rather, they behave like grammatical elements, which makes it challenging to distinguish them using part-of-speech tags. Once the poems were divided into 320-word consecutive segments, all other words except nouns, verbs, and adjectives were removed in accordance with the part-of-speech tags. The part-of-speech tags for the analysis are displayed in Table 3 with indications from the UCREL CLAWS5 Tagset (<https://ucrel.lancs.ac.uk/claws5tags.html>) in brackets.

Table 3. Part-of-speech tag list to be analyzed

Nouns	NN0	[Common noun, neutral for number]
	NN1	[Singular common noun]
	NN2	[Plural common noun]
Verbs	VVB	[The finite base form of lexical verbs]
	VVD	[The past tense form of lexical verbs]
	VVG	[The -ing form of lexical verbs]
	VVI	[The past participle form of lexical verbs]
	VVN	[The infinitive form of lexical verbs]
	VVZ	[The -s form of lexical verbs]
Adjectives	AJ0	[Adjective (general or positive)]
	AJC	[Comparative adjective]
	AJS	[Superlative adjective]

The Machine Learning for Language Toolkit was utilized to apply LDA to the segments (McCallum, 2002). The number of topics was determined to be 20, based on the findings from earlier experimental trials, which ranged from 10 to 100. The optimal number of topics was determined by evaluating the LDA results and the close reading of the poems. After reducing the number of candidate topics ranged from 18 to 30, the author repeated the LDA process for each number of topics. Generally, the LDA results vary with each run; however, the 20-topic configuration produced consistent outcomes because the terms assigned to each topic varied minimally. Therefore, the author of this study decided that 20 is the most fitting number of topics for further analysis.

3. Results of the LDA

The LDA output results are discussed below; some results point out topics for

further discussion. The observation of the outcome provides the answer to the first research question of this article: 1. Can LDA detect the differences in poetic diction between works by Alfred and Charles? Table 4 presents a subset of the LDA output results, including topic numbers, alpha values, and the most salient keywords associated with each topic. The keywords are arranged in descending order of weight from top left to bottom right. The universality of each topic is indicated by the alpha values: a lower alpha value denotes that the topic appears in fewer segments, while a higher alpha value indicates that the topic appears more prominently across multiple segments in the corpus.

The heatmap representation (Figure 1) provides a visual illustration of the representativeness of the topics, with colors indicating the degree of representation. The darker the cells, the higher the density of the given topic in the poems, and the whiter cells represent a significantly lower topic density. The 20 topics were arranged in a vertical sequence, and the groups of poems (segments) were aligned in a horizontal sequence. The 995 segments for 525 poems have been classified into 12 groups, as it was not feasible to display all 525 works in a single heatmap, given the limitations of a standard sheet of paper. The titles of each group denote the following: A27, Alfred's poems from *Poems by Two Brothers*; ac27, poems with doubtful authors (Alfred or Charles) from *Poems by Two Brothers*; Aup, Alfred's poems included in *Poems by Two Brothers* 1893 edition but not in the 1827 edition; C27, Charles's poems from *Poems by Two Brothers*; T20s, Alfred's poems, which were written in the 1820s but remained unpublished and were not included in the 1827 or 1893 editions of *Poems by Two Brothers*; T30s–T90s, Alfred's poems published or written in the stated range of years. The vertical line titles show the topic numbers with the two most prominent keywords of each topic connected by an underscore (e.g., merry_bride for Topic 0).

As illustrated in the heatmap in Figure 1, the cells representing the most general topics, Topics 5, 12, and 13, are predominantly represented by darker colors. The differentiation between the three topics is based on the frequency of appearance of specific groups. The cells for Topic 5 are observed to be colored darker in the groups of poems published/written by Alfred during the period from the 1840s to the 1890s (T40s–T90s). Topics 12 and 13 exhibit a darker hue in the groups comprising the authors' early works, which were published/written from the 1820s to the 1830s. While Topic 12 demonstrates a darker coloration in the groups where Charles's name is

Table 4. Output result of LDA (Topic number, alpha values, and top 20 keywords)

Topic	Alpha values	Top 20 keywords
0	0.0483	merry_a, bride_n, bridegroom_n, days_n, shake_v, bone_n, cow_n, borne_v, plains_n, praise_n, milking_v, dance_n, whisper_v, green_a, silent_a, grow_v, sits_v, cause_n, goes_v, month_n
1	0.1147	throne_n, war_n, high_a, king_n, woe_n, glorious_a, pride_n, sword_n, proud_a, trumpet_n, fame_n, fiery_a, glory_n, fire_n, course_n, bow_n, battle_n, earthly_a, rise_v, strength_n
2	0.4006	came_v, said_v, heard_v, saw_v, went_v, man_n, knew_v, hand_n, made_v, fell_v, took_v, stood_v, ring_n, dead_a, left_v, spoke_v, turn'd_v, ran_v, isle_n, look'd_v
3	0.0774	death_n, men_n, glory_n, ship_n, battle_n, fight_v, hill_n, sea_n, roof_n, rode_v, arm_v, die_v, right_n, banner_n, fight_n, wives_n, gallant_a, bold_a, devil_n, fought_v
4	0.0613	follow_v, whirl_v, sun_n, winds_n, morning_n, hide_v, forgotten_v, science_n, song_n, slope_n, ends_n, wrong_a, grown_v, woo_v, fine_a, youth_n, earth_n, jewel_n, faint_v, won_v
5	0.7318	heart_n, life_n, love_n, let_v, old_a, know_v, day_n, little_a, love_v, world_n, mother_n, child_n, friend_n, night_n, gone_v, go_v, dead_a, happy_a, good_a, true_a
6	0.0670	pride_n, bones_n, city_n, vale_n, midnight_n, solemn_a, echoing_a, form_n, dark_a, hangs_v, broad_a, cloudy_a, holy_a, grave_n, sombre_a, varied_a, steadfast_a, shades_n, valley_n, branches_n
7	0.0502	time_n, golden_a, good_a, prime_n, old_a, goose_n, honour_n, goodly_a, reason_n, great_a, rhyme_n, worthy_a, side_n, place_n, teach_v, immortal_a, pleasure_n, harder_a, weather_n, stream'd_v
8	0.2223	flowers_n, golden_a, sweet_a, green_a, air_n, year_n, sing_v, song_n, fair_a, summer_n, happy_a, birds_n, flower_n, spring_n, music_n, brook_n, fresh_a, young_a, tree_n, river_n
9	0.0706	glowing_a, god_n, realms_n, fire_n, secret_a, lyre_n, boundless_a, lute_n, lustre_n, maze_n, chords_n, countless_a, bard_n, harp_n, lay_n, roses_n, give_v, sway_n, magic_a, reign_n
10	0.4263	man_n, things_n, power_n, human_a, time_n, nature_n, men_n, life_n, world_n, faith_n, truth_n, mind_n, soul_n, words_n, age_n, great_a, wise_a, make_v, years_n, times_n
11	0.1368	land_n, great_a, queen_n, people_n, men_n, let_v, name_n, freedom_n, king_n, free_a, war_n, voice_n, kings_n, hearts_n, isles_n, cause_n, ancient_a, health_n, friends_n, sea_n
12	0.6745	eyes_n, heart_n, love_n, life_n, death_n, soul_n, sweet_a, tears_n, face_n, light_n, hope_n, fair_a, spirit_n, eye_n, see_v, low_a, mind_n, place_n, full_a, joy_n
13	0.6726	light_n, night_n, earth_n, day_n, sun_n, deep_a, heaven_n, dark_a, bright_a, voice_n, sky_n, sound_n, moon_n, white_a, sea_n, wind_n, stars_n, high_a, eyes_n, cloud_n
14	0.0775	gate_n, lime_n, garden_n, oak_n, seed_n, city_n, boughs_n, eddies_n, maiden_a, folded_a, boat_n, broad_a, windy_a, fern_n, read_v, bridge_n, feed_v, rock_n, beech_n, farm_n
15	0.0677	gods_n, priest_n, bread_n, god_n, fire_n, cross_n, holy_a, flesh_n, saints_n, saved_v, mercy_n, prayer_n, sin_n, mountain_n, saint_n, plague_n, people_n, word_n, hymns_n, leper_n
16	0.0486	let_v, ring_v, form_n, riflemen_n, storm_n, rave_n, cup_n, look_v, rose_n, grave_n, lisette_n, dainty_a, ready_a, green_n, wine_n, fill_v, grave_a, folds_v, order_n, warm_a
17	0.0734	eye_n, ringleet_n, bright_a, charm_n, soft_a, touch_n, sure_a, charms_n, wing_n, view_n, gay_a, gale_n, ecstasy_n, passing_a, scene_n, beauties_n, shrine_n, appear_v, dye_n, virtue_n
18	0.0518	brows_n, seemed_v, smile_n, stood_v, verge_n, smiling_v, flitting_v, gleam_n, bridge_n, clearer_a, frown_n, constant_a, sense_n, floating_a, inward_a, rays_n, wandering_v, idol_n, lamb_n, solid_a
19	0.0476	poet_n, art_n, popular_a, sake_n, muses_n, laurel_n, claim_n, fame_n, gave_v, sow_n, bailiff_n, price_n, college_n, prate_v, fire_n, friends_n, wrong_a, line_n, days_n, grant_v

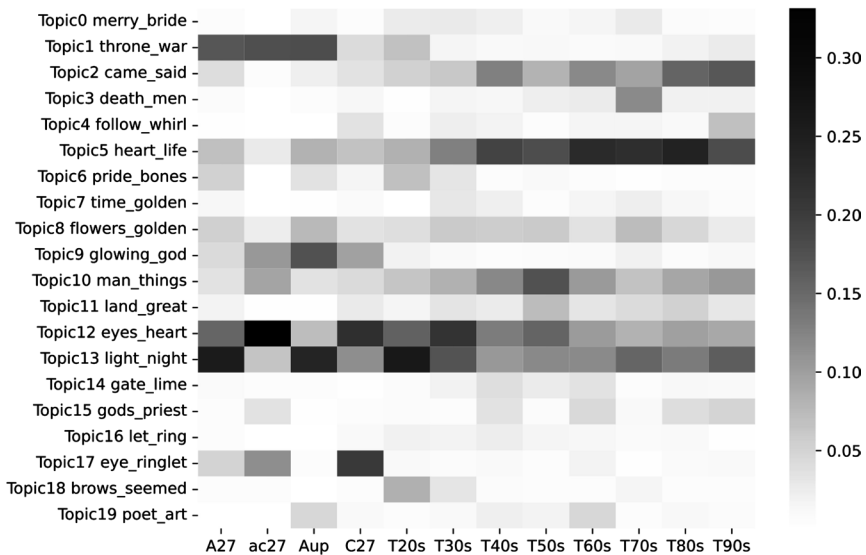


Figure 1. Heatmap of topic densities in clonological categories

assigned (namely, ac27 and C27), Topic 13 also comprises dark cells in

Figures 2 and 3 present the top 50 mean density poems for the two topics, which are identified in the preceding paragraph as Topics 17 and 1, respectively. The heatmap in Figure 1 depicts the average densities of the topics for the 12 groups. Figures 2 and 3, in contrast, represent more specific densities for the poems. Given that some poems were split into multiple segments during the LDA and the densities were assigned per segment, the density values for several segments for one poem were calculated and averaged to plot the bar charts.

Topic 17 is a prevalent topic in Charles's poems, as illustrated in Figure 2. A total of 33 poems are assigned to Charles, but 17 Alfred's works, published in various years, are also included in the top 50 poems containing Topic 17. Within 17 Alfred's poems of Topic 17, 11 poems were from the collection of *Poems by Two Brothers* (1827). Of the 11 poems, nine were marked as written by Alfred and two poems were doubtful but assumed to be by Alfred. The six poems discussed above by Ricks are included in the top 50 poems of Topic 1's density (Figure 3). In addition to the six poems, 24 poems were drawn from *Poems by Two Brothers* (1827). Four works were written/published

in the 1820s by Alfred, but they are not included in the collection *Poems by Two Brothers*. Excluding the doubtful author works “The Deity” (ac17_ITY), “The Dying Christian” (ac27_IAN), and “Switzerland” (a27_AND), the 41 poems of Topic 1 are

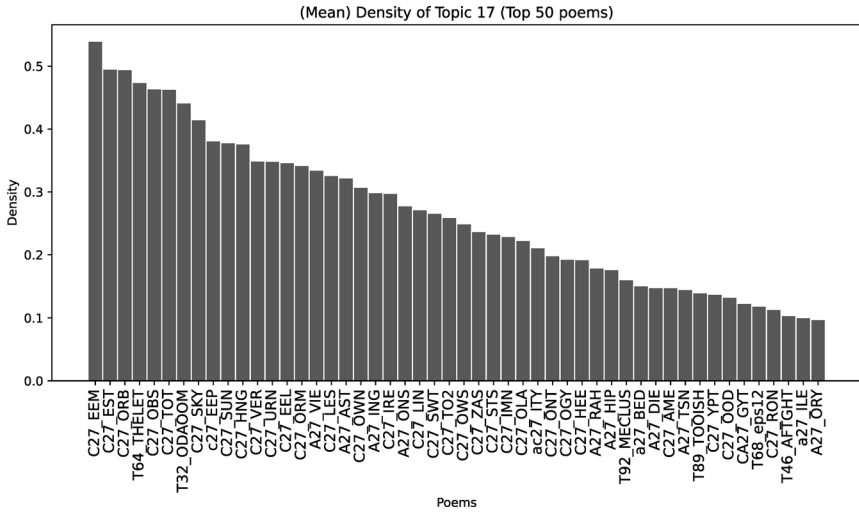


Figure 2. Bar plot of the (mean) density of Topic 17

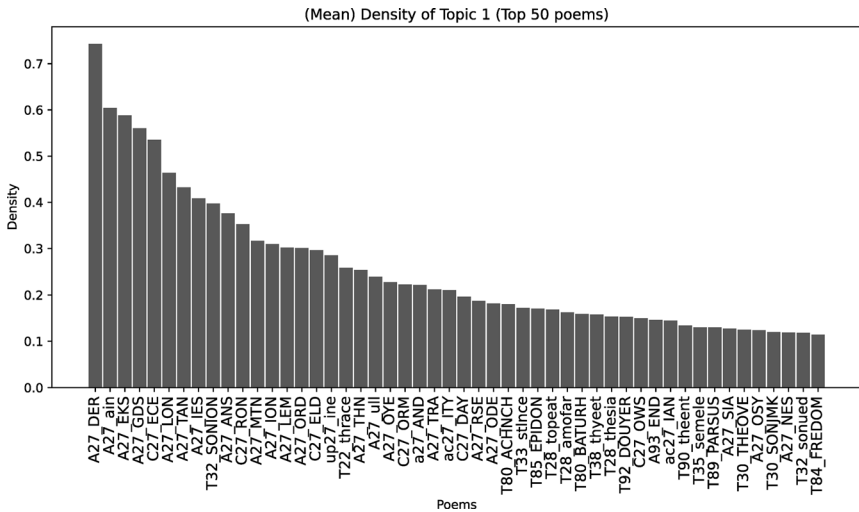


Figure 3. Bar plot of the (mean) density of Topic 1

Alfred's poems.

The results of the LDA analysis indicate that Topic 17 is heavily represented in Charles's poems, whereas Topic 1 is notably present in Alfred's poems from 1827. Furthermore, Topic 9 also emerged in Charles's oeuvre, albeit with a lower frequency than Topic 17. It should be noted that these topics did not exclusively manifest in the works of just one of the brothers. With that acknowledged, the LDA outcomes indicated the tendencies of topics for both Alfred and Charles. In this regard, the initial research question, "Can LDA detect the differences in poetic diction between Alfred's and Charles's works?" can be answered in the affirmative. The following section further narrows the discussion and provides meticulous observation of the two most prevalent topics mentioned in this section. The discussion section of this paper thus attempts to elucidate answers to the second research question, namely, "If LDA detects differences, what characteristics do the two authors exhibit?"

4. Discussion

In this section, the author undertakes an observation and discussion of two topics that were previously identified: Topics 17 and 1. Topic 17 was featured primarily in Charles's poems, whereas Topic 1 is particularly evident in Alfred's poems from 1827. This section is divided into two sections, with each section addressing a specific topic: Topic 17 is discussed in section 4.1 and Topic 1 in section 4.2.

4.1 Topic 17

Topic 17 was a recurring topic in Charles's poems. The terms within the top 20 keywords of Topic 17 are predominantly employed to convey romantic sentiments, feelings toward others (particularly women), and references to women's physical appearance, including the keywords *eye*, *bright*, *charm*, *touch*, *charms*, *ecstasy*, and *beauties*. The noun *ecstasy* is the thirteenth keyword of Topic 17, and is exclusive to the poems of *Poems by Two Brothers*. Alfred used it on two occasions, and Charles utilized it eight times. The poems exude profound romantic passion and ardor, as evidenced by the following lines (boldface added; from this point onward, the use of bold text in quotations will indicate that the referenced word has been assigned to the topic under discussion): "To gaze on thee is **ecstasy**, /Is **ecstasy** — but pain." ("Oh

were This Heart of Hardest Steel,” ll. 25–26; C27_EEL) and “Why did I burn with feverish **ecstasy**, /Stung with her scorn, and ravish’d with her praise?” (“The Slighted Lover,” ll. 7–8; C27_VER). The fervor of the language and the lines themselves were uniquely present in the 1827 collection, evoking a sense of youthful vigor and association. The poems’ content prompted Hallam Tennyson, a son of Alfred Tennyson to say, “As an outburst of youthful poetic enthusiasm, the book is not wanting in interest and a certain charm, although full of the boyish imitation of other poets” (Tennyson H., 1897, p. 22).

Despite the fact that the words assigned to Topic 17 are seen throughout Charles’s works, the results revealed that Alfred’s poem from 1864, “The Ringlet” (T64_THE-LET), also ranked within the fourth density of Topic 17 (Figure 2). In “The Ringlet,” the term *ringlet* appears with notable frequency. The 11th line of the poem states, “My **ringlet**, my **ringlet**,” and the 25th line continues with the repetition of “O **Ringlet**, O **Ringlet**.” There are 10 instances in the poem where similar lines repeat the term *ringlet* twice in a line bringing the total occurrences of the word to 20. The term *ringlet* is the second keyword within Topic 17, yet it does not appear in Charles’s works. A total of 24 instances of the term *ringlet* were identified within the corpus. Of these, 20 occurred within the poem “The Ringlet,” while the remaining four were distributed across “The Talking Oak” (1842), *In Memoriam A.H.H.* (1850), and “The Ring” (1889) written by Alfred.

The term *ringlet*’s occurrences were all assigned to Topic 17, although its use is not exclusive to Charles. The third keyword of Topic 17, *bright*, exhibits a distinctive pattern from Topic 17. The total number of occurrences of the word *bright* in the entire corpus is 172, and its frequency it appears in Topic 17 is 19. The 153 instances were allocated to Topics 12 or 13, which are more universal topics than Topic 17, as indicated by their alpha values. Of the 19 appearances of the term *bright* in Topic 17, 13 were found in Charles’s works and six in Alfred’s. When using LDA, it is not uncommon for the same term to be sorted into different topics. This is due to the fact that LDA utilizes the concept of a “bag of words” concept, whereby the probability of co-occurrence is analyzed within a given segment. If the bags of words exhibit disparate patterns of co-occurrence, the words within the bags can be assigned to distinct topics. The following excerpts illustrate the poems where *bright* appears, where one is assigned to Topic 12 and the other to Topic 17. Alfred’s poem, “The Grave of a Sui-

cide” (A27_IDE), is presented on the left, and Charles’s poem, “The Slighted Lover” (C27_VER), is presented on the right.

HARK! how the gale, in mournful notes and stern,
Sighs thro’ yon grove of aged oaks, that wave
(While down these solitary walks I turn)
Their mingled branches o’er yon lonely grave!

...

For thou, wed to misery from the womb —
Scarce one bright scene thy night of darkness knew!

(“The Grave of a Suicide”: ll. 1–4, 11–12)

I LOVED a woman, and too fondly thought
The vows she made were constant and sincere;
But now, alas! in agony am taught,
That she is faithless — I no longer dear!

Why was I frenzied when her **bright** black eye,
With ray pernicious, flash’d upon my gaze?

(“The Slighted Lover”: ll. 1–6)

The term *bright*, which is underlined in Alfred’s “The Grave of a Suicide,” has been assigned to Topic 12, while the *bright* in Charles’s “The Slighted Lover,” which appears in bold, has been assigned to Topic 17. Although the excerpts do not display the entirety of each poem’s lines, the differences between the two can be discerned. While Alfred’s work demonstrates a sense of lamentation pertaining to the life of a person (*yon*), Charles’s poem illustrates the sentiment of remorse experienced by the individual (*I*). The term *bright* is employed in two distinct ways in the two poems. In Alfred’s poem, it is used to signify both light in the darkness and hope in the context of a person’s miserable life. In Charles’s poem, however, it is used simply to modify the description of a lady’s (*her*) “black eye.”

The singular word *eye* was the most significant keyword of Topic 17; it was assigned to Topic 17 in the fifth line of the excerpt of Charles above, where the third

keyword *bright* modifies. The frequency of the singular word *eye* was 118 across 80 works, while the frequency of plural *eyes* was 260 across 138 works within the 525 works composing the corpus. The discrepancy between the frequency of the singular and plural forms of the word *eye* does not invalidate the intuitive assumption that the plural form is more prevalent. This is because the human body has two eyes, and the *eye* is defined as “one of the two parts of the body” (s.v. *eye*, *n.* 1.: *Longman Dictionary of Contemporary English* (2014); underline added by author). Consequently, when the organ is referenced in language describing body parts, the plural form *eyes* is often employed. A concordance line of *eyes* in the corpus is shown in Figure 4 as an example of plural *eyes* usages. The singular form *eye* is not precluded, however, and there may be motivations or reasons for the distinct usages of plural and singular forms of the word in both Charles’s and Alfred’s works.

Kwic - 260 found in 138 files	File Name
1 evel In the new birth of those imperial eyes .	T30_owaep.txt
2 ne Returning with hot cheek and kindled eyes .	T72_ALEDER.txt
3 y blood, Light of the Light within mine eyes .	T42_lfood.txt
4 ce that froze my lay Be melted by thine eyes!	C27_IGN.txt
5 ht, but as lissome as a hazel wand; her eyes a bashful azure, and her hair In gloss	T55_THEOOK.txt
6 ound and like a pear In growing, modest eyes , a hand a foot Lessening in perfect cad	T42_WALAIL.txt
7 ing to pass. In much amaze he stared On eyes a bashful azure, and on hair In gloss a	T55_THEOOK.txt
8 malice now – But often in the sidelong eyes a gleam of all things ill – It is not L	T85_THEGHT.txt
9 ng from the ledge: And when I raised my eyes , above They met with two so full and br	T33_THETER.txt
10 not pen a Sonnet lest the scorner With eyes acute my glaring faults pursue, And hem	up27_idd.txt
11 Of the maiden, that I lost, In my inner eyes again , Lest my heart be overborne, By t	T37_THEAUD.txt
12 n of the dead no more – my child! Thine eyes Again were human-godlike, and the Sun B	T89_DEMONE.txt
13 t into my heart, and begun to darken my eyes . Ah – you, that have lived so soft, wha	T80_RIZPAH.txt
14 ood name.’ And he turn’d, and I saw his eyes all wet, in the sweet moonshine: Sweet	T62_THEHER.txt
15 n girl alone, Serene with argent-lidded eyes Amorous , and lashes like to rays Of dar	T30_RECHTS.txt
16 o the melody, till they flew, Hair, and eyes , and limbs, and faces, Twisted hard in	T42_THESIN.txt
17 he blessed skies For which I live? black eyes , and brown and blue; I hold them all mo	T31_SONNGS.txt
18 won the heart, Lovely with bright black eyes and long black hair And lips which move	T75_SENNCE.txt
19 dal bower: On me she bends her blissful eyes And then on thee; they meet thy look An	T50_INMAHH.txt
20 nd a maiden tender, shy, With fair blue eyes and winning sweet, And longed to kiss h	T46_became.txt
21 hath drawn the frozen rain From my cold eyes and melted it again.	T30_SONWOE.txt
22 Strain the hot spheres of his convulsed eyes , And in his writhings awful hues begin	T30_LOVOVE.txt
23 climb, Floats from his sick and filmed eyes , And something in the darkness draws Hi	T30_SUPELF.txt
24 Of hairy strength, and white and garish eyes , And silent intertwisted thunderbolts,	T28_armdon.txt
25 solace as he may. Menoeceus, thou hast eyes , and I can hear Too plainly what full t	T85_TIRIAS.txt

Figure 4. Concordance lines of “eyes” from the corpus

Focusing on the singular word *eye* assigned to Topic 17, the term was observed a total of 26 times across 18 works. Of the 26 instances, 17 were found in Charles’s poems, nine were present in Alfred’s poems. Among the 17 instances of the *eye* in Charles’s works, seven were observed between the lines shown in following (1)–(7), while the remaining ten instances were found at the end of the lines. In excerpts (5) and (7), *eye*-s are referred to as a singular entity due to the grammatical requirements of the language. In line (5), the preceding adjective *each* demands a singular noun form. In

From the sad spot, and in mine **eye**

The full warm tear-drop burn’d.

(“A Sister Sweet Endearing Name” (C27_AME): ll. 17–20; bold added)

But winter came — its varied **dye**

Each morn grew fainter to mine **eye**;

Till, with’ring, it was bright no more,

Nor bloom’d as it was wont before:

(“Still Mute and Motionless She Lies” (C27_STS): ll. 13–16; bold added)

Given that the pronunciation of *eye* is comprises of a single diphthong, /ai/, it can be inferred that the entirety of the word *eye* itself, or the entire sound of the word /ai/, represents the target for rhyming with another word. Departing from the Topic 17 elements and contemplating Charles’s oeuvre in a more comprehensive manner, however, an intriguing suggestion emerges. The act of rhyming is typically understood to entail the utilization of identical or analogous vocal elements at the end, beginning, and/or middle of poetic lines. The same or similar sounds are based on vowels, and it is not necessary for the consonants preceding or following the vowel to be identical. Additionally, as the term *similar* indicates, the vowel (and consonant) sound(s) need not be an exact match. With this established, in Charles’s poems, the exact match of a vowel and the subsequent consonant(s) frequently occurs: for example, in *pow’r/flow’r*, *roll/pole*, *fire/ire* (“In Summer when All Nature Glows” (C27_OWS): ll. 19–24); in *stage/age*, *view/woo*, *awake/take*, *steals/heels/reveals* (“Still Mute and Motionless She Lies”: ll. 1–9). As previously stated, Charles’s “Still Mute and Motionless She Lies” provides an illustrative example of *eye/dye* rhyming. It commences with a rhyme involving the plural *eyes*:

STILL, mute, and motionless she lies,

The mist of death has veil’d her eyes.

And is that bright-red lip so pale,

Whose hue was freshen’d by a gale

More sweet than summer e’er could bring

To fan her flowers with balmy wing!

(“Still Mute and Motionless She Lies”: ll. 1–6; underline added)

A comparison of the *eyes* in the second line and *eye* in the 14th line (as seen in

the previous excerpt) of the “Still Mute and Motionless She Lies” reveals a distinct contrast between the two lines. The contrast hinges on the pronoun used to describe the *eye/eyes*: the genitive case third-person pronoun *her* or the first-person possessive pronoun *mine*. Upon expanding our “eye” to include the anteroposterior lines, terms rhyming with *eye/eyes* are observed to differ between the lines. Given that the rhymes in Charles’s work frequently align with both vowel and subsequent consonant sounds, it is reasonable to hypothesize that the use of singular or plural nouns in analogous positions may be influenced by rhyme or sound structures.

Similar to *eye/eyes*, the singular and plural forms of *charm* were among the top 20 keywords of Topic 17. The singular form of *charm* was found three, nine, and eight times in Alfred’s works from 1827, in Alfred’s works from the 1830s to the 1890s, and in Charles’s works, respectively. The plural form of *charms* appeared two, three, and nine times in Alfred’s works from 1827, in Alfred’s works from the 1830s to the 1890s, and in Charles’s works. In Charles’s oeuvre, there is only one instance of *charms* occurring at the end of a line, where it rhymes with *warms* (“Imagination” (C27_IMN): ll.15–16), and otherwise, the word is found in the middle of the lines. Conversely, the three instances of Alfred’s poems are positioned at the end of the lines, and all three uses of the word *charms* rhyme with *arms*. The aforementioned examples of the distinctions between the singular and plural forms of *charm* indicate that there is not a single, straightforward reason or motivation underlying the differences in usage. Nevertheless, the observations on singular and plural differences in one topic indicated the possibility that sound preferences might be a contributing factor in the alteration of their forms. It is regrettable that LDA is unsuitable for the analysis of sound and grammatical elements. It is therefore not possible to conclude that LDA has identified the rhyming preferences of the author(s). To ascertain the rhyme and/or sound structures and preferences of the both Charles and Alfred, further analyses employing optimal methods are required in future studies.

4.2 Topic 1

Topic 1 appeared significantly in the *Poems by Two Brothers*. Of the top 50 density poems in Figure 3, 30 were included in the 1827 collection. In addition to the 30 poems from 1827, five other poems were written by Alfred in the 1820s. Additionally, the 43 poems among the top 50 poems of Topic 1 are also Alfred’s poems. Topic 1 can

therefore be considered a topic that primarily represents Alfred’s poems, particularly those written during his early career. Unlike the poems of Topic 17, the poems in which Topic 1 frequently appears tend to primarily address masculinity or substances, evoking images of men and scenes in which men are often depicted. Of the top 20 keywords of Topic 1, the fourth keyword, *king*, was directly related to the concept of a male crown. Other keywords, including *throne*, *fame*, and *bow*, are associated with notions of nationhood and royal authority. The keywords *war*, *sword*, *fire*, *battle*, and *strength* are linked to both nations and masculinity, as historically, men have been the ones to serve their nations or crowns. Wars or battles are often initiated for the purpose of protecting or expanding a nation, region, or diadem. Further observation revealed additional relationships between keywords. The terms *glorious*, *pride*, *proud*, *fame*, *glory*, and *trumpet* are strongly associated with the keywords *war* and *battle*. These associations are evident in various poems by Alfred that appear in the 1827 collection. The following quotations are “The High Priest to Alexander” (A27_DER) and “Exhortation to the Greeks” (A27_EKS), the first and third density poems of Topic 1.

Go forth, thou man of force!

The world is all thine own;

Before thy **dreadful course**

Shall **totter** every **throne**.

Let India’s **jewels** glow

Upon thy **diadem**:

Go, forth to **conquest go**,

But **spare** Jerusalem.

For the **God** of **gods**, which **liveth**

Through all **eternity**,

’Tis He alone which **giveth**

And **taketh victory**:

(“The High Priest to Alexander”: ll. 1–12)

AROUSE thee, O Greece! and **remember** the day,

When the millions of Xerxes were **quell’d** on their

way!

Arouse thee, O Greece! let the **pride** of thy name

Awake in thy bosom the light of thy **fame!**

...

Remember each day, when, in **battle array**,

From the fountain of **glory** how largely ye **drunk!**

For there is not aught that a freeman can **fear**,

As the **fetters** of **insult**, the name of a **slave**;

And there is not a voice to a nation so dear,

As the **war-song** of **freedom** that calls on the **brave**.

(“Exhortation to the Greeks”: ll. 1–4, 21–26)

As evidenced by the aforementioned poems, the top 20 keywords of Topic 1 can be discerned not only in isolation but also in conjunction with their synonyms and related terms assigned to the topic. In “The High Priest to Alexander,” the term *diadem* is associated with the crown of a nation. Additionally in this same poem, the term *victory* is related to the concepts of *glory*, *war*, and *battle*, which are among the top 20 keywords of Topic 1. In “Exhortation to the Greeks,” *quell’d*, *slave*, *war-song*, and *freedom* are correlated with the concept of *battle*. It is apparent that the locations of these battles and wars were not necessarily within the boundaries of the United Kingdom, as illustrated by references to *Jerusalem* and *Greece* in “The High Priest to Alexander” and “Exhortation to the Greeks.” Furthermore, the top 20 keywords of Topic 1, as well as the poems themselves, demonstrate a sense of masculinity or vigor, despite the paucity of references to individuals in the poems.

The following excerpt is from “Written During the Convulsions in Spain” (A27_ain), the second highest density poem of Topic 1, written by Alfred. In this poem, the top 20 keywords, as well as the words related to the keywords, such as *arm*, *combat*, and *fight*, were observed as the terms of Topic 1. In addition, the term *heroes*, connoting masculinity, was identified.

Strong be their **arm** in **war**,

Brilliant their **glory**’s star,

Fierce be their **valour** and **fearful** their name!

...

Where are thine **heroes** hid?

Arm them for **combat** and shout, ‘To the **fight!**’

Shake the **throne** of thy Lord

To its **base** with their **sword**,

So, on to the **combat**, and God **help** the right!

(“Written During the Convulsions in Spain”: ll. 16–18, 32–36; bold added)

Masculinity also emerges in Charles’s poems, yet these evince disparate emotional nuances compared to those expressed by Alfred. The following quotes from Charles’s poems are not as vigorous as Alfred’s ones, whereas several Topic 1 keywords are employed in Charles’s poems. In the excerpt of “On the Death of Lord Byron” (C27_ROM), the terms *hero*, *career*, *blaze*, and *fame*, are designated as the Topic 1 keywords. In this poem, the singular *hero* refers to George Gordon Byron (1788–1824), who is regarded as a representative poet of the Romantic era. Despite the absence of any explicit references to warfare or combat, the poem’s principal subject is a male figure. While previous literature, such as by Shaw (1973) and Thomas (2019), has indicated similarities in the poetic styles of Alfred and other Romantic poets (for example Percy Bysshe Shelley, John Keats, and William Wordsworth), there is a paucity of lamentations in Alfred’s poetry regarding the loss of other Romantic poets and a dearth of commentary on the works of other Romantic poets. In “On the Death of Lord Byron,” however, Charles expresses great fervor in his lamentation of the loss of Lord Byron. Although the existence of “On the Death of Lord Byron” does not directly refute the notion that Alfred held other poets in high regard, it does imply that Alfred’s respect for and engagement with other poets may not be as profound and deep as Charles’s. The enthusiasm respectively displayed by Alfred and Charles diverged during their adolescence, however. Charles’s poem “The Battle-field” (C27_ELD) depicts a scene of battle or war but does not include the terms *battle* or *war*. Despite this, terms related to warfare or battle are ascribed to Topic 1, including *chaos*, *contest*, *madden*, *trumpet*, *barbarous*, *bray*, and *cannon*. Likewise, the use of the term *heroes* in “The Battle-field” suggests that the poem also extols masculinity.

THE **hero** and the bard is gone!

His bright **career** on earth is done,

Where with a comet’s **blaze** he shone.

...

Was Byron's hope — was Byron's aim:
With ready heart and hand he came;
But perish'd in that path of **fame!**

(“On the Death of Lord Byron”: ll. 1–3, 37–39)

THE heat and the **chaos** of **contest** are o'er,
To mingle no longer — to **madden** no more:
And the cold forms of **heroes** are **stretch'd** on the
plain;
Those lips cannot breathe thro' the **trumpet** again!

...

I — heard, oh! I heard, when, with **barbarous bray**,
They leapt from the mouth of the **cannon** away;

(“The Battle-field”: ll. 1–4, 9–10)

While a handful of Charles's poems are included in the top 50 poems of Topic 1, the number of Alfred's poems in the topic was significantly greater. It can be concluded that LDA identified the predominantly male elements, masculinity, and enthusiasm for and in aspects of battle in both Charles's and Alfred's poems in Topic 1, appearing more frequently however in works by Alfred.

5. Conclusion

This study employed the quantitative approach LDA to identify the characteristic diction of Alfred Tennyson and his brother, Charles, in their first publication, *Poems by Two Brothers* (1827). The LDA outcomes indicated that two topics, Topics 17 and 1, were particularly prevalent in the collection. Furthermore, Topic 17 was identified as a more prominent feature in the poems of Charles, while Topic 1 was observed to be a significant element in Alfred's poems. Topic 17 was found to represent terms associated with romantic sentiments directed toward women and descriptions of their physical appearances. In contrast, Topic 1 represented lexical items associated with masculinity, enthusiasm, and battles. The distinction between the two topics suggests that there are

differences in the vocabulary and pattern of expression used by Alfred and Charles. Topics 17 and 1 yielded responses to our initial research question, “Can LDA detect the differences in poetic diction between Alfred’s and Charles’ works?” Upon examination of these topics, we were also able to ascertain answers to our secondary research question, “If LDA detects differences, what characteristics do the two authors exhibit?”

Previous authorship attribution studies have employed neither the LDA nor the analysis of content words, which were the focus of this study. The objective of the present study was not to ascertain the efficacy of LDA in authorship attribution but to identify internal evidence of the distinguishing characteristics of the two brothers. Indeed, the results indicated that the differences between the authors could not be fully classified. The findings of this study indicate the limitations of LDA in terms of achieving complete certainty regarding authorship attribution. However, the accuracy of author estimation can be further enhanced by combining the results of other function words and examining content words using LDA in the context of lyrical poetry studies. This is because quantitative lyrical poetry studies are confronted with the challenge of handling shorter and smaller data than prose text and gaining reasonable data size and results.

Section 4.1, above, posits the potential influence of sound preferences based on the observation of the themes of Topic 17. It would be beneficial for future studies to consider the stylistic features of the poems, including function words and rhymes, to gain insight into the authorship of the poems. Nevertheless, the distinctive diction indicated in this article will constitute an element of internal evidence. Concerning diction, future studies are needed to observe a greater number of poems and topics in order to identify other possible features of each poet. The integration of internal and external evidence, in addition to qualitative and quantitative approaches, will surely further advance the authorship attribution and the study of Alfred’s and Charles’s poems.

Note

1 This study employs the reprinted edition of *Poems by Two Brothers* (1827), printed in London by W. Simpkin and R. Marshall, and J. and J. Jackson. The reprinted edition was published by Thomas Y. Crowell in New York, though the precise date of publication is not indicated.

Acknowledgment

This study is supported by a grant from the Japan Society for the Promotion of Science (JSPS) KAKENHI program (grant number 24K22483). Expressions of gratitude are extended to the anonymous reviewers who have provided constructive feedback, leading to significant enhancements in the quality of this article. It should be noted that the responsibility for any residual errors remains with the author.

Bibliography

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Brimley, J. R. (1972). *Tennyson & His Poetry*. New York: AMS Press.
- Fujita, I. (2020). “He triumphs; maybe, we shall stand alone”: Using correspondence analysis to investigate modal adverbs in Tennyson’s poems. *Japanese Association of Digital Humanities 2020*, 21–26.
- Fujita, I. (2022). On segment size in poetry analysis using the latent dirichlet allocation method. *Digital Humanities*, 3, 3–15.
- 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.
- Jockers, M. L. (2014). *Text Analysis with R for Students of Literature*. Heidelberg, New York, Dordrecht, London: Springer Cham.
- Leech, G., & Smith, N. (2000, March 17). *Manual to accompany: The British National Corpus (Version 2) with Improved Word-class Tagging*. Retrieved November 25th, 2024 from https://ucrel.lancs.ac.uk/bnc2/bnc2postag_manual.htm
- Longman Dictionary of Contemporary English. (2014). Retrieved November 25th from Longman Dictionary of Contemporary English: <https://www.ldoceonline.com/>
- McCallum, A. K. (2002). *MALLET: A machine learning for language toolkit*. Retrieved November 25th from <http://mallet.cs.umass.edu>
- Paden, W. D. (1964). The Tennysons’ *poems by two brothers* (1827) reconsidered. *The Library*, 5, 147–161.
- Ricks, C. (1987). *The Poems of Tennyson* (2nd ed.) (Vol. III). London: Longman.
- Shaw, D. W. (1973). *Tennyson’s Style*. Ithaca: Cornell University Press.
- Tabata, T. (2018). Mapping Dickens’s novels in a network of words, topics, and texts. *Text mining and Digital Humanities, 2017*, 47–58.
- Tennyson, A.; Tennyson, C. (n.d.). *Poems by Two Brothers*. New York: Thomas Y. Crowell.
- Tennyson, A.; Tennyson, C. (1893). *Poems by Two Brothers* (2nd ed.). London: Macmillan.
- Tennyson, A. (2013). *Delphi Poets Series Alfred, Lord Tennyson*. East Sussex: Delphi.
- Tennyson, H. (1897). *Alfred Lord Tennyson: A Memoir by his Son*. London: Macmillan.

Thomas, J. (2019). *Tennyson Echoing Wordsworth*. Edinburgh: Edinburgh University Press.
UCREL CLAWS5 Tagset. (n.d.). Retrieved November 25th, 2024 from CLAWS part-of-speech
tagger for English: <https://ucrel.lancs.ac.uk/claws5tags.html>

(Iku Fujita, currently at Kyushu University; formerly at Hijiyama University,
E-mail: tnnysn.annie.if@gmail.com)