

Title	LDAトピックモデルによるIPAテキスト分析の試み : アルフレッド・テニスの韻文を用いて
Author(s)	藤田, 郁
Citation	言語文化共同研究プロジェクト. 2022, 2021, p. 15-38
Version Type	VoR
URL	https://doi.org/10.18910/88358
rights	
Note	

Osaka University Knowledge Archive : OUKA

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

Osaka University

LDA トピックモデルによる IPA テキスト分析の試み

—アルフレッド・テニソンの韻文を用いて—

藤田 郁

大阪大学言語文化研究科

〒560-0043 豊中市待兼山町 1-8

Email: u256780k@ecs.osaka-u.ac.jp

概要 本稿は、Latent Dirichlet Allocation (LDA) トピックモデリングが詩の韻律の分析にも応用できるかどうかを検証することを目的としている。LDA はテキストに隠された意味構造を検出する文体分析手法であり、現在では主に散文テキスト分析へ応用する研究が多い。LDA を用いた詩に関する研究は比較的少ない中で、Navarro-Colorado (2018) は、スペイン語で書かれた詩においても LDA で韻を検出できる可能性を示唆した。しかし、英語では、綴り字と発音・音は必ずしも一致せず、また音節や音素の単位で押韻される例も数多く見受けられる。そのため、綴り字のままの英文テキストに LDA を適用することは、テキスト中の韻を検出するための最適な方法とはいえないと考えた。本研究では、英語の綴り字と音の不一致を考慮し、国際音声記号 (IPA) で書かれたテキストと、IPA で書かれた語を母音と子音に分割したテキストの2種類のテキストを分析することにより、IPA テキストへの LDA 適用の可能性、及び IPA テキストや音単位のテキストを分析する際の問題点を提示する。

キーワード トピックモデル, LDA, 国際音声字母, アルフレッド・テニソン, 詩

An Exploratory Investigation of IPA Texts with LDA:

Using Alfred Tennyson's Poems

Iku Fujita

Graduate School of Language and Culture, Osaka University

1-8 Machikaneyama-cho, Toyonaka, Osaka, 560-0043 Japan

Abstract This paper aims to investigate whether Latent Dirichlet Allocation (LDA) topic modelling can also detect the rhymes in poetry. LDA is a stylometric method that detects hidden semantic structures in texts and is now used primarily to examine prose texts. Among a relatively small number of studies on poems using LDA, Navarro-Colorado (2018) suggested that LDA can also detect rhymes in Spanish-written poetry corpus. However, speaking of the English language, the spellings and the pronunciations/sounds do not always work correspond. Additionally, numerous rhymes can be found structured by the units of syllables or phonemes. Thus, applying LDA on raw English texts may not be the optimised method for detecting rhymes in texts. Considering the dissidence between spelling and sound of English, this study analyses two text-types, written in International Phonetic Alphabets (IPA) and a text of which all words in IPA are divided into vowels and consonants. Through the analysis and discussion

of the emerging results of LDA, this paper provides the possibility of employing LDA on IPA texts as well as issues of analysing IPA and more small-unit texts.

Keywords Topic Model, latent Dirichlet allocation, International Phonetic Alphabets, Alfred Tennyson, Poetry

1. はじめに

現在、デジタル人文学・計量的手法は人文学研究、文学研究及びイギリス文学研究に大きく貢献している。文学研究における研究対象はもっぱら散文作品が多いが、韻文作品も例外ではなく、この手法は現在さまざまな側面から詩の研究に新たな知見をもたらしている (Tartakovsky and Shen, 2019; Nakao and Jimura, 2016 等)。

計量的手法の中でも、トピックモデルは、テキストマイニングの分野において将来性のある手法であるとされている (Meeks & Weingart, 2012)。本論で用いるトピックモデリング手法 latent Dirichlet allocation (Blei et al.; 2003; 以下 LDA) は、特に多くの散文作品研究に用いられている (Tabata, 2017; Kuroda, 2017; Kiyama, 2018; Huang, 2020a, b)。しかし、トピックモデルの韻文作品への応用は、Rhody (2012), Navarro-Colorado (2018), Henrichs (2019), Okabe (2019) 等を除き、散文作品と比べ未だ少なく、また LDA を国際音声字母 (International Phonetic Alphabets: 以下 IPA) の分析に応用した例は、筆者の知る限りではまだなされていない。

LDA とは、テキストコーパスにおける隠れた意味構造を明らかにする機械学習手法であり、テキスト内での (語の) 共起傾向を基に、語をグループ分けする教師なし手法であり (Tabata, 2017: 52), 意味を持つ最小単位である語を分析するために通例用いられる。筆者もこれまでの研究では、引喩の中にある Tennyson 作品の特徴を明らかにすべく、語に着目した分析を行なっている。スペイン語で書かれたソネット¹を LDA を用いて分析と考察を行なった Navarro-Colorado (2018) は、以下のように述べ、LDA がソネットに適用された場合に押韻する語も検出できることを示唆している。

[T]he LDA-topic is used as a poetic motif, considering motif here as a set of words that usually appear together. Two kinds of motifs have been defined: phonetic motifs and semantic motifs. Phonetic or “rhyme” motifs include Topics whose keywords have similar soundS [sic], that is, they rhyme.

Navarro-Colorado (2018: 11)

しかし、押韻は語単位で行われるよりもむしろ語よりも小さい音節あるいは音素単位で行われるものであり、加えて英語の発音は綴り字と一致しないものも多く、綴り字は同じでも同じ発音となるとは限らない (例: -oo-, **cook** /u/; **choose** /u:/)。また、「韻を踏む」という条件は、同じ語を使用する必要はなく、語やその意味、文法的機能などに関係なく、以下 (1) の例のように、母音や子音が同じ・似ている音で構成されている語を使用することも多い。英語詩において押韻が行

¹ 十四行詩とも呼ばれる。14 行から成る定型詩。

われる単位（音節や音素）と LDA が分析する単位（語）が異なっている以上， Navarro-Colorado (2018) の押韻する語も検出できるという指摘は，少なくとも英語詩には当てはまらない可能性があると考えられる。LDA の結果が押韻を検知したように見えたとしても，それは副産物としての結果であり，語単位での LDA を用いた分析は，押韻を検知させることを主目的にすることは妥当ではないと考える。

(1) () 内は IPA 表記，[] 内は押韻の種類を表示。（太字下線は筆者による）

THEY rose to where their sovran eagle sails ,	(/s <u>eilz</u> /)	[1]
They kept their faith, their freedom, on the height ,	(/h <u>ait</u> /)	[2]
Chaste, frugal, savage, arm'd by day and night	(/n <u>ait</u> /)	[2]
Against the Turk; whose inroad nowhere scales	(/s <u>keilz</u> /)	[1]
Their headlong passes, but his footstep fails ,	(/f <u>eilz</u> /)	[1]
And red with blood the Crescent reels from fight	(/f <u>ait</u> /)	[2]
Before their dauntless hundreds, in prone flight	(/fl <u>ait</u> /)	[2]
By thousands down the crags and thro' the vales .	(/v <u>eilz</u> /)	[1]
O smallest among peoples! rough rock- throne	(/θr <u>əʊn</u> /)	[3]
Of Freedom! warriors beating back the swarm	(/sw <u>ɔ:m</u> /)	[4]
Of Turkish Islam for five hundred years ,	(/j <u>iəz</u> /)	[5]
Great Tsernogora! never since thine own	(/ɔ <u>ʊn</u> /)	[3]
Black ridges drew the cloud and brake the storm	(/st <u>ɔ:m</u> /)	[4]
Has breathed a race of mightier mountaineers.	(/n <u>iəz</u> /)	[5]

“Montenegro” (1877)

押韻している箇所をグルーピングする目的に特化するのであれば，国際音声字母 (International Phonetic Alphabets; 以下 IPA) に変換したものをを用いることで，音を統一された基準で視覚的に捉えられる「形」で示すことができると同時に，homophone (異形同音異義語; read と reed, right と write, I と eye 等) のように音が同じにもかかわらず綴り字の違いにより別のグループに分類される問題を解決でき，トピックモデルの（特に）韻文作品への応用可能性が広がるのではないかと考えた。一方で，トピックモデルは先にも述べたように意味を持つ最小単位の語を分析の対象とし，コーパス内の文書に跨る意味的つながりを見つけることができるアルゴリズムである。ここでの「意味」とは，語の持つ辞書的な「意味」ではなく，またアルゴリズムの内部にセマンティックな辞書を有しているわけでもいため，IPA で書かれたテキストを LDA で分析することにより，見出す「意味」を「押韻（要素）」とすることは可能だと考える。しかし前述の通り，LDA を IPA テキストに応用した例はまだなく，意味を持たない単位である音節や音素に焦点を当てた場合にトピックモデルが適切な分析手法となり得るか，また音の背後にある要素＝韻を検知できるか否かを検討する必要がある。本稿では，まず raw テキストと IPA テキストを比較す

ることで、IPA に変換したのみのテキストに LDA を実行することによってどのような違いが得られるか検討し、更に IPA テキストを子音と母音とで分割した場合に得られる結果を検討する。

2. データ及び方法

2.1. データ

Delphi Poet Series 20, Complete Works of Alfred, Lord Tennyson に収録されている韻文作品から、作品毎のトークン数が 100~199 の作品 116 編を本稿の分析対象とする。トークン数が 100~199 である作品群は、Tennyson の作品のおよそ 4 分の 1 を占め、最も作品数が多い群である。加えて、本作品群にはソネットをはじめ、(脚) 韻を踏んでいる作品が多く、IPA テキストの分析結果を LDA がどのように示すか、押韻等の要素を検知できているか否かを検証するには、押韻の(少ない自由詩等を用いるよりも適切であると考えた。116 編の作品は、本文のみがテキストファイルに取められており、テキストファイルのタイトルに短縮した作品の題名と発表年が付してある。これらのファイルを一括して raw テキストとし、IPA 表記ファイル作成の元データとするとともに、分析結果の比較対象として用いる。

IPA 表記への変換は、オンラインで使用可能である toPhonetics (<https://tophonetics.com>) というコンバータを利用した。IPA の表記は、イギリス英語とアメリカ英語が選択できるほか、IPA に変換された文字列とともに変換前の英文を併記させることも可能であり、また、and の発音における /ænd/ と /ənd/ に見られるアッシュと (æ) シュワ (ə) の差のような強形と弱形表記を選択することができる。本稿で用いるデータは、全てイギリス英語、IPA テキストのみ表示、強形の設定で変換した。尚、コンバータの辞書にない単語は、IPA に変換されることなく入力されたままの文字列で出力される。この変換されなかった文字列は、本稿の分析では英語表記のまま分析にかけることとした。上記の手順を踏まえ、ファイルごとに変換作業を行い、出力された結果を新たなテキストファイル (raw ファイルと同様にファイル名を付与済み) に書き込んだものを IPA テキストとして分析対象データとした。

3 節で考察を行うデータは、先に用意した IPA テキストの内容を子音と母音で分割した IPA sep (arated) データである。テキストエディタに正規表現を入力し、連続する子音 (例: /sk/ (/a:sk/)) /母音字 (例: /əʊ/ (/səʊbə/)) は分割せず、あくまで子音と母音が連続している箇所を分割するよう設定した。そのため、この分割は音節や音素と同等の音の単位を示していない。これら分割した IPA 表記は、本項では「音」あるいは「(音の) 要素」として扱う。また、本稿では半母音 /j/ は子音として扱っている。分割に際して、IPA テキストに使用されていた強勢記号をはじめ記号類は、母音の長音記号のみを残して削除した。また、分割する際にその音 (群) が語頭にあったか、語中、語末にあったかを示すために、ギリシャ文字 φ (phi; ファイ) を用い、語頭の場合には文字の後方のみ、語中の場合には文字列の前後に、語末の場合には文字の前方にのみ φ を付与した。φ を用いたのは、このマークをハイフンなどの記号にしてしまうと、LDA を実行する際に語の一部として認識されず、自動的に半角スペースに置換されてしまうため、英語及び英語の IPA 発音表記では用いられず、また視覚的にも判別しやすい文字を付与するためである。以

下 (2) に、各テキスト内のテキストデータの一部を例として示す。

(2)

raw テキスト： DEAR, near and true — no truer Time himself
 IPA テキスト： dɪə, niəɹ ænd tru: — nəʊ 'tru:ə taɪm hɪm' self
 IPA sep テキスト： dɪə fɪə nɪə fɪəf ɪə æf ɪnd tɹɪə fu: nɪə fəʊ tɹɪə fu:ə tɪə fəɪf ɪm hɪə fɪə
 fɪmsɪə fɛf ɪlf

“A Dedication” (1862: l. 1)

各ファイルのトークン数は以下表1の通りである。尚、アポストロフィ (') 及びハイフン (-) では語を分けないものとした。IPA sep テキストでは、子音と母音を分割した後のスペースで区切られている文字列 (例：fɪəf) を一語としてカウントしている。Raw テキストと IPA テキストにおける異なり語数の差は30語となっているが、これは同音異綴りのタイプ違い (例：our, hour; /aʊə/) が90例、異音同綴り (例：after; /ɑ:ftə(r)/) が60例あり、その差分が異なり語数の差に反映されている。なお、同音異綴りには、いわゆる同音異義語 (例：made, maid; /meɪd/) のほか、所有格 -'s と複数形 -s (例：heart's, hears; /hɑ:ts/), 過去形・過去分詞 -ed と -'d (例：looked, look'd; /lʊkt/) の接尾辞による違いがみられた。異音同綴りでは、冠詞の a/the の強形 (/eɪ/, /ði/) 弱形 (/ə/, /ðə/) の違い以外は全て語末の /t/ 音の有無によるものであった。

表 1 使用したデータ毎のトークン数と異なり語数

	raw テキスト	IPA テキスト	IPA sep テキスト
総トークン	15,528	15,528	45,065
異なり語数	3,651	3,621	718

2.2. データの実行及びトピック数の決定方法

LDA を実行するにあたり、MALLET 2.0.8 をインストールし、分析を行う。LDA 実行の際に任意のストップワードを指定することができるが、本分析では一切の語を削除せずに分析を行なった。また、LDA では、研究者 (実行者) が予めトピックの数を設定する仕様となっている。本稿では、可能な限り恣意性を排除することを目的として、perplexity (モデルの予測精度の評価指標) と coherence (トピックの品質を測る評価指標) をテキスト種類毎に計算した結果から、適切なトピック数を10種類まで絞り込み、LDA を実行して得られる α 値 (以下3節で詳述) と分析結果を比較検討し、トピック数を決定した。分析及び考察に用いる最適なトピックの数は、raw テキストと IPA テキストは10、IPA sep テキストは17をそれぞれ最適なトピック数とした。

3. 分析及び考察

本節では、raw テキスト, IPA テキスト, IPA sep テキストの3 テキストに対し、それぞれ LDA を実行し、得られた結果を提示する。

3.1. Raw テキストと IPA テキスト

3.1.1 Raw テキスト

本節ではまず、raw テキストの結果を提示する。以下表2に示すのは、raw テキストへの LDA 実行結果のうち、各トピックの α 値と各トピックにおける最も顕著なキーワード (Keys) 上位 20 語である。トピックは表の左から右に向かい α 値の降順で並んでおり、Keys は上から下に向かい、顕著度合いが高いものから低いものへ、降順で並んでいる。例えば、最も α 値が高いトピックは、本出力結果では Topic 6 ($\alpha = 1.88294$) であり、最も α 値が低いトピックは Topic 3 ($\alpha = 0.07916$) である。また、Topic 6 において、最も顕著な語の上位 1 位は the であり、20 位の語は as である。後の 3.1.2 節及び 3.2 節の表 4, 6 についても同様の表示をする。

表 2 LDA による各トピックの α 値及び上位 20 語 (raw テキスト)

Topics	Topic 6	Topic 4	Topic 5	Topic 2	Topic 1	Topic 7	Topic 0	Topic 8	Topic 9	Topic 3
α value	1.88294	0.21179	0.16916	0.16534	0.16431	0.14020	0.11504	0.09728	0.09695	0.07916
Keys (1-20)	the	i	thy	her	you	we	rosalind	our	riflemen	o'
	and	my	thou	she	your	our	valley	england	form	each
	of	me	thee	on	i	its	nothing	land	far	both
	to	thro'	oh	light	it	ye	falls	lord	storm	they
	in	brother	thine	sleeps	my	why	stream	their	your	together
	a	only	art	scorn	men	must	through	every	ready	side
	that	tho'	yet	form	again	ah	until	we	words	two
	with	eyes	heart	cold	me	gaze	along	freedom	he	other
	for	lilian	beware	saw	know	e'en	strange	farewell	nor	than
	is	hand	ere	ever	three	woe	full	god	city	beauty
	all	heart's	stay	nor	is	tear	living	through	sound	tears
	but	morning	wilt	had	room	tho'	awearry	part	against	look
	so	love	truth	with	hangs	us	bold	macready	do	an'
	from	with	year	rest	him	welaway	even	blood	trumpet	ralph
	not	thee	kiss	doth	little	grief	winter	thine	now	perfect
	his	mind	pass	year	should	bright	spring	thee	him	stars
	when	loved	canst	not	here	ecstasy	years	years	charm	over
	on	tis	sleep	down	read	soon	mourn	thought	thunders	a'
	be	dare	alas	hair	loved	joys	about	where'er	angel	luve
	as	sad	hope	after	leaf	glow	earth	french	speed	sighs

α 値は、高ければ高いほど、そのトピックが該当コーパスに多く含まれていることを示す値である。一方で α 値が非常に低い場合には、そのトピックがごく少数のテキストにしか現れず、場合によってはある 1 つのテキストにしか現れていないことが多い。現に、表 2 では、最も α 値が高い Topic 6 の Keys は 20 語全て機能語であり、また表 4 に示す Raw テキストにおける高頻度語上位 20 位にそのほとんどが含まれており、テキスト内で広く使われている語であることが分かる。反対に、 α 値が最も低い Topic 3 には、特定の詩にしか使われていない固有名詞 (ralph) が

上位語に含まれており、このトピックが少数の詩にしか現れていないことが分かる。

3.1.2. IPA テキスト

次に、表 3 に示すのは IPA テキストに LDA を実行した出力結果、各トピックとその α 値、キーワード (Keys) 上位 20 語である。この IPA テキストのトピック数も、先の raw テキストと同じく 10 である。

表 3 Raw テキストにおける高頻度語上位 20 位

(下線は表 2 Topic6 の上位 20 語に含まれていることを示す)

Rank	Word	Freq	Rank	Word	Freq
1	<u>the</u>	993	11	her	134
2	<u>and</u>	615	12	s	127
3	<u>of</u>	451	13	<u>is</u>	120
4	<u>to</u>	256	14	<u>for</u>	110
5	<u>in</u>	247	15	my	109
6	<u>a</u>	233	16	<u>all</u>	103
7	<u>that</u>	196	17	thy	101
8	i	174	18	<u>but</u>	96
9	<u>with</u>	165	19	<u>not</u>	94
10	d	145	20	<u>from</u>	91
			20	<u>so</u>	91

表 4 LDA による各トピックの α 値及び各トピック上位 20 語 (IPA テキスト)

Topics	Topic 3	Topic 8	Topic 4	Topic 9	Topic 0	Topic 7	Topic 1	Topic 6	Topic 5	Topic 2
α value	1.94883	0.31232	0.19183	0.16167	0.14637	0.13617	0.13258	0.10989	0.10312	0.06632
Keys (1-20)	ðə	aɪ	ðaɪ	hɜ:	wi:	wɪl	fɑ:	ju:	lɔ:d	bi:
	ænd	maɪ	ðəʊ	ʃi:	'aʊə	rosalind	əʊ	jo:	gɒd	ðeɪ
	ʊv	mi:	ði:	hɜ:r	ɪts	wɪnd	ən	əʊld	ðeə	rɪfləmən
	tu:	ɪt	ðəm	nɒt	wɑ:	'i:vən	hi:	hʌm	bɪ'fɔ:	fɔ:m
	m	səʊ	nɒt	ɒn	mʌst	stri:m	hʌm	hæd	bəʊnd	ɑ:
	ə	ə	θru:	fɔ:m	wəʊ	daʊn	ðæn	gru:	stɑ:z	stɔ:m
	ðæt	ə'gen	ɑ:t	sli:ps	vɒt	'fɛmɔ:z	lɒk	jo:r	ænd	ðeər
	wɪð	əʊ	jet	lɪliən	geɪz	'væli	səɪz	hɜ:d	blæd	'redi
	ɪz	lʌvd	wi:	nɔ:	deɪ	'fri:dəm	bɪ'ni:θ	fli:t	rælf	laɪk
	ɔ:l	nju:	nɔ:	ɑ:	sti:l	ə'lɒŋ	ə'weɪ	pɑ:st	kɪŋ	hɑ:ts
	fɹɒm	'brʌðə	'aʊə	hɑ:t	gləʊ	lænd	hæŋz	bɪ'fɔ:	tɔ:k	səɪd
	bʌt	wɪ:p	bɪ'weə	kæən	həʊp	'nʌθŋ	'hevɪli	bɪ'kɔ:z	streɪm	wi:
	wɛn	'lɪtl	lʌv	laɪt	həʊ	fɔ:lz	'swi:tə	træst	sɛd	deɪ
	ɒn	ju:	sɪns	ə'bʌv	bʌt	vɔɪs	ə'baʊt	lɒkt	ðeəz	tə'geðə
	hɪz	'bjʊ:ti	'feə'wel	ɪt'self	'ɛkstəsi	sæt	eɪ	mi:	bɪ:st	'ɪŋglənd
	fɔ:	mɔ:	steɪ	'vɜ:tju:	bɪ'fɔ:r	'ʃaɪld	luve	hʌ:m	ðɛm	weləweɪ
	æz	'əʊnli	hɑ:t	fɪ:t	grɪ:f	streɪndʒ	wɜ:dz	leɪt	grɛt	rɪ'vjuəz
	nəʊ	hænd	aɪ	'pɜ:fɪkt	teə	'mɪdl	teəz	gest	hu:	'ɪndələnt
	ðeə	sʌf	macready	hɑ:f	ʃi:	ə'wɪəri	səʊnd	ə'fɑ:	'ɑ:ftər	'pi:pl
	bɑ:	fi:l	vɔɪs	hænd	reɪz	fɹeɪʃ	hu:z	ræɪmz	'ʃu:li	fɹeɪnʃ

表 2 に挙げた raw テキストの結果と比較すると、トピックの番号や α 値は異なるものの、各トピックの上位語に共通している点が多く見られる。上位 20 語に最も共通する語が多いのは、raw テキストの Topic 6 と IPA テキストの Topic 3 であり、20 語中 17 語が同じ語である。最も共通する語が少ないトピックは、Topic 1 (raw) と Topic 6 (IPA), Topic 8 (raw) と Topic 5 (IPA) の 4 語である。全体では、各テキスト全 200 キーワードのうち、重複して現れた語は、その 6 割にあたる 120 語であった。

Raw テキストの Topic 4 と IPA テキストの Topic 8 は、共にキーワードの上位 3 語が一人称単数代名詞の主格 I, 所有格 my, 目的格 me となっているほか、Topic 5 (raw) と Topic 4 (IPA) では二人称単数代名詞及び二人称所有代名詞 thy, thou, thee, thine, Topic 2 (raw) と Topic 9 (IPA) では三人称単数女性代名詞 her, she, そして Topic 1 (raw) と Topic 6 (IPA) では二人称代名詞の you, your がトピック毎に分類されている。Tennyson の作品を LDA で分析した際に、人称代名詞が各トピックに人称や単複、性差によって分かれ、作品毎の登場人物や語り手の違い、登場人物間の関係や作品のテーマ分類される点は、Fujita (2021) で得られた示唆と共通する点である。

以下図 1, 2 に示すのは、LDA の分析から得られる構成 (composition) データを用いて描画したヒートマップである。このヒートマップは、どのトピック (縦軸) がどの作品 (横軸) に顕著に現れているかを色の濃淡で示すものであり、セルが濃い黒であればあるほど、そのトピックはセルが交差する軸に位置する作品に顕著なトピックであると言える。このヒートマップから、例えば共通して一人称単数代名詞がキーワードに含まれている Topic 4 (Raw) と Topic 8 (IPA) が顕著に現れる、セルが他のセルより濃い黒で示されている作品を比較する。Topic 4 (Raw) は、主に 24 作品に顕著に現れるトピックであり、Topic 8 (IPA) は 31 作品に顕著に現れていることが分かる。この二つのトピックが顕著である作品数は異なっているものの、それぞれに共通する作品は、以下表 5 に示す Topic 4 (Raw) と Topic 8 (IPA) に顕著な作品のリストにおいて網掛け太字で示している 19 作品である。

表 5 Topic 4 (Raw) と Topic 8 (IPA) に顕著な作品のリスト

Raw: Topic 4				IPA: Topic 8					
1	Ancrntc31	13	MyRslnd32	1	Ancrntc31	13	OHWRTHS27	25	ThSptfl68
2	ANDSKYW27	14	PrftryP80	2	ANDSKYW27	14	OLDG89	26	ThThrst89
3	APLLNSR27	15	Sonnt-Bu32	3	APLLNSR27	15	Songvry30	27	TISTHVC27
4	ASSTRSW27	16	SonntC130	4	ASSTRSW27	16	Sonnt-Bu32	28	To32
5	Bynvltn89	17	Sonnthb32	5	Bynvltn89	17	SonntC130	29	ToLnhst53
6	Crssngt89	18	THBRDSF27	6	DobtndP92	18	Sonnthb32	30	Toswhnw32
7	DobtndP92	19	THDYNGC27	7	Herftnw50	19	ThBggrM42	31	YESTHRB27
8	Herftnw50	20	Thrrthr32	8	InthGrd74	20	THDYNGC27		
9	Liln30	21	THSLGHT27	9	Liln30	21	THGRVFS27		
10	LovndSr30	22	ThSlrBy62	10	MYBRTHR27	22	Thrrthr32		
11	Memywnf32	23	TISTHVC27	11	MyRslnd32	23	THSLGHT27		
12	MYBRTHR27	24	ToLnhst53	12	ODrlngR32	24	ThSlrBy62		

以上のように、raw テキスト及び IPA テキストへの LDA 実行結果は比較的類似している。ま

た、Topic 2 (IPA) のキーワードに含まれる /fɔ:m/ (form) と /stɔ:m/ (storm) は、押韻を LDA が検知した例のようにみえる。この 2 語は、raw テキストの Topic 9 の上位 2 語でもある。Topic 2 (IPA) が顕著に現れる 4 作品、MiltN-H62, Dulms30, RiflMnF92, NatlSn30 のうち、単語が使用されているのは以下 (3) に挙げる RiflMnF92 (“Riflemen Form”, 1892) のみである。作品原文をみると、確かに storm と form は脚韻である。しかし、storm と form 以外にも韻を踏んでいる語、音が複数確認できるが、これらの語は LDA の結果には反映されておらず、押韻要素が検出されていると断定できるものではない。これらのことから、IPA テキストにおいても raw テキストにおいても語の持つ意味とその共起傾向から、トピックが分類されていることが分かる。加えて、raw, IPA 各テキストの総語数 15,528 に対して、述べ異なり語数の差は 150 であり、その量、また種類を考慮すると、raw テキストと IPA テキストとでは、大きな差があるとは言えず、本稿で扱うデータにおいては LDA を実行するテキストを IPA 表記にただけでは、音韻要素を LDA が検知できるか否かは不明である。そこで、IPA テキストを母音字と子音字に分割したデータ (IPA sep テキスト) を用いて再度 LDA を実行した結果を、次の 3.2 節で提示する。

(3) () 内は (行末 1 語の IPA 表記; 押韻している箇所の IPA 表記)

THERE is a sound of thunder afar,
 Storm in the South that darkens the *day!* (/deɪ/; /eɪ/)
 Storm of battle and thunder of war! (/wɔ:/; /ɔ:/)
 Well if it do not roll our *way.* (/weɪ/; /eɪ/)
 Storm, Storm, Riflemen **form!** (/fɔ:m/; /ɔ:(m)/)
 Ready, be ready against the **storm!** (/stɔ:m/; /ɔ:(m)/)
 Riflemen, Riflemen, Riflemen **form!** (/fɔ:m/; /ɔ:(m)/)
 Be not deaf to the sound that warns, (/wɔ:nz/; /ɔ:nz/)
 Be not gull'd by a despot's *plea!* (/pli:/; /i:/)
 Are figs of thistles? or grapes of thorns? (/θɔ:nz/; /ɔ:nz/)
 How can a despot feel with the *Free?* (/fri:/; /i:/)

“Riflemen Form” (1892; ll. 1–11; 飾り字は筆者による)

3.2. IPA sep テキスト

本節では、IPA sep テキストの結果を提示する。IPA sep テキストに LDA を実行する際のトピック数は、17 とした。以下表 6 に示すのは、各 17 トピックの α 値と各トピックにおける最も顕著なキーワード (Keys) 上位 20 語である。トピックは表の左上段から右下段へ α 値の降順で並んでおり、Keys は上から下に向かい、顕著度合いが高いものから低いものへ、降順で並んでいる。

表 6 において最も α 値が高いトピックは、Topic 3 ($\alpha = 6.10462$) であり、最も α 値が低いトピックは Topic 6 ($\alpha = 0.09079$) である。最も α 値が高く、複数の作品に現れる Topic 3 のキーワードをみると、raw テキストの高頻度語 (表 7) を構成する音が多く含まれていることが分かる。

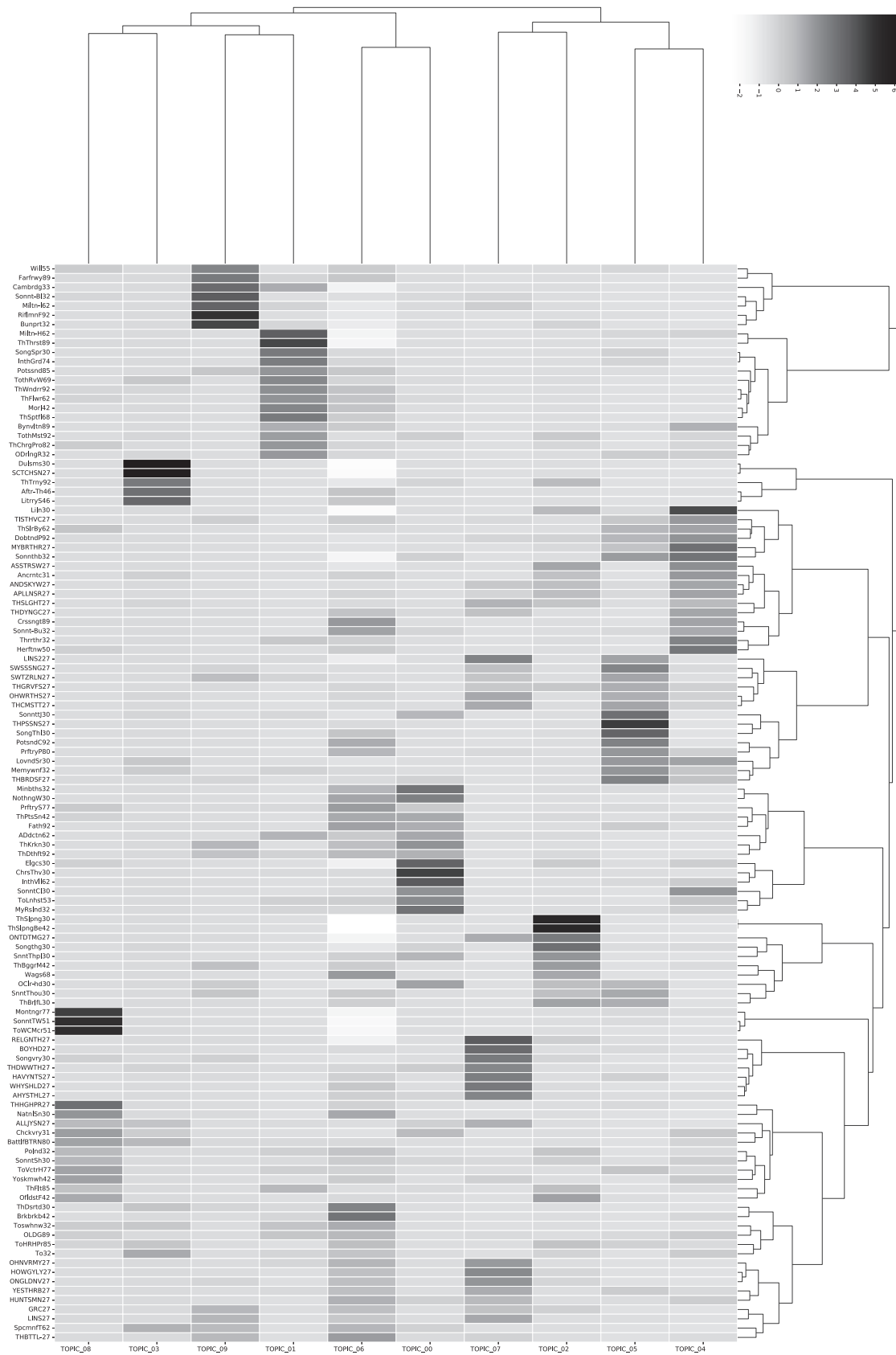


図 1 ヒートマップ (raw テキスト)

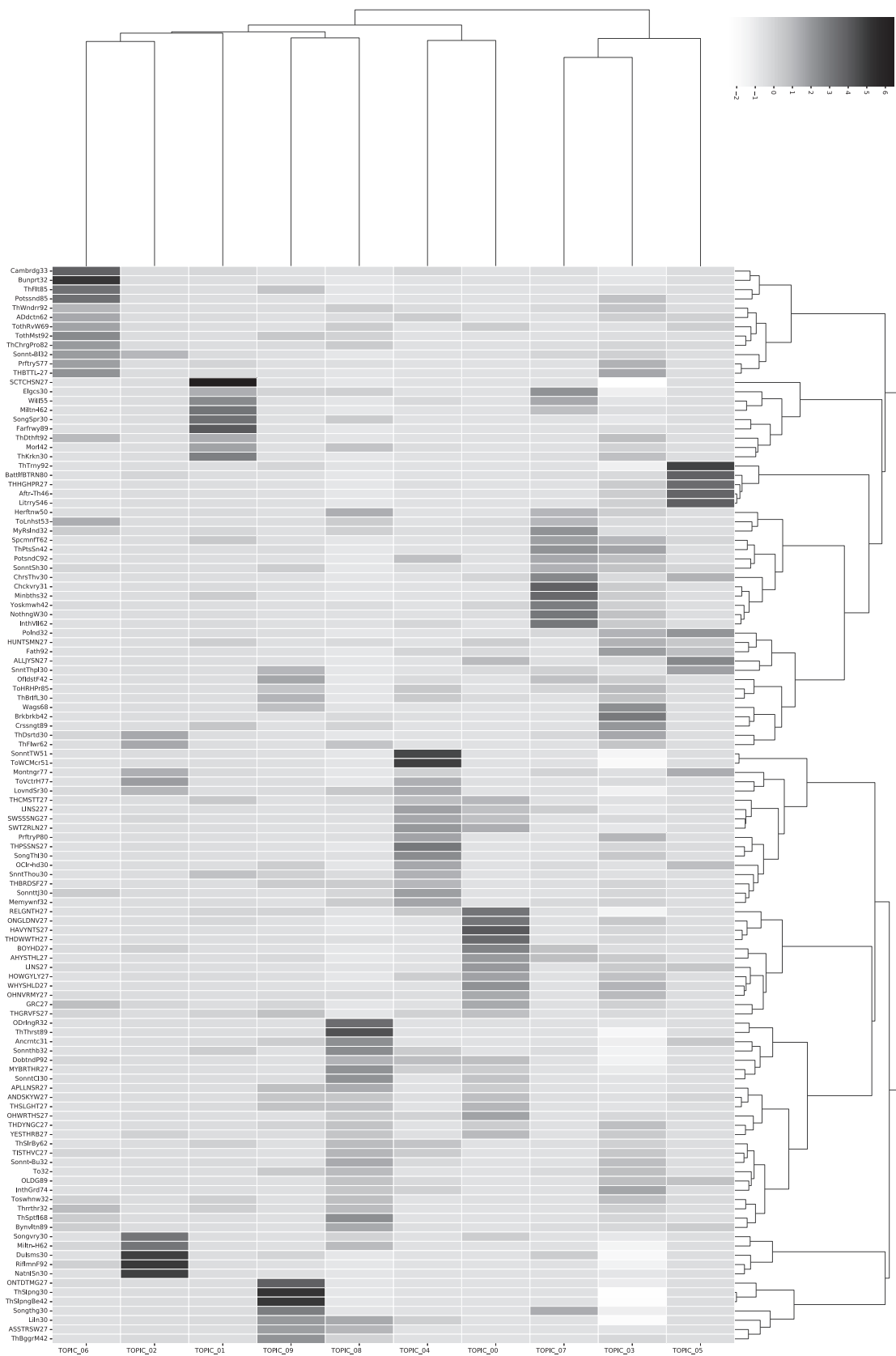


図 2 ヒートマップ (IPA テキスト)

しかし、高頻度語を構成する音の全てが Topic 3 に現れているわけではない。表 7 の頻度第 8 位にある i は、IPA 表記の場合は eye と同形となり、さらに頻度が増えるため、 α 値が高いトピックに分類されることを予想したが、前後に子音のつかない単独 ai は、 α 値が比較的低い Topic 11 にのみ現れている。このように、複数のトピックで、元の raw テキスト及び IPA テキストにおける語の頻度や、LDA の出力結果から推測できそうな分類が見られる一方で、IPA テキストに LDA を実行した際には見られなかった音のばらつきもある。

表 6 LDA による各トピックの α 値及び各トピック上位 20 語 (IPA sep テキスト)

Topics	Topic 3	Topic 10	Topic 14	Topic 12	Topic 4	Topic 13	Topic 2	Topic 5	Topic 15
Alpha	6.10462	3.61320	3.44713	2.93176	2.54886	1.86466	1.72640	1.42596	1.09778
Key (1-20)	ɸnd	ɸz	ɸiɸ	ðɸ	ɸl	ɸəɸ	ðɸ	ɸiɸ	ɸiɸ
	æɸ	iɸ	wɸ	ɸə	ðɸ	ɸiɸ	ɸə	ɸlɸ	wɸ
	ɸn	lɸ	ɸi:ɸ	ɸv	ɸai	ɸrɸ	bɸ	ɸəʊɸ	bɸ
	ɸeɸ	ɸaiɸ	ɸz	ɸɸ	ɸt	ɸs	ɸaiɸ	ɸi	ɸə
	ɸæɸ	ɸd	sɸ	ɸn	ɸæɸ	ɸi	ɸeiɸ	ɸɸɸ	ɸŋ
	ɸm	ɸəʊ	ɸt	ɸɸɸ	ɸn	ɸɸ	ɸə	ɸfɸ	ɸð
	ɸeiɸ	mɸ	ɸəɸ	ɸəʊɸ	mɸ	ɸv	ɸəʊɸ	ɸntɸ	nɸ
	hɸ	sɸ	rɸ	dɸ	tɸ	ɸk	ɸvɸ	ɸnɸ	ɸlz
	ɸu:	ɸa:ɸ	ɸŋ	frɸ	kɸ	ɸnɸ	ɸæɸ	ɸ:ɸ	ɸwɸ
	ɸɔ:ɸ	hɸ	ɸi	pɸ	ɸeɸ	ɸu:ɸ	ɸeɸ	mɸ	strɸ
	fɸ	ɸt	ɸfʃ	ɸtɸ	ɸs	vɸ	ɸnz	spɸ	ɸz:ɸ
	ɸt	bɸ	brɸ	ɸld	ɸi:	glɸ	ɸei	ɸpɸ	ɸvɸ
	wɸ	nɸ	iɸ	trɸ	ɸd	ɸiæɸ	ɸaʊɸ	flɸ	θrɸ
	ɸl	ɸʌɸ	blɸ	gɸ	ʃɸ	ɸtɸ	ɸaiə	a:ɸ	æɸ
	ɸi:	ɸei	ɸd	ɸmz	fɸ	ɸzɸ	ɸts	ʌɸ	ɸɔ:ɸ
	ɸʌɸ	ɸi:ɸ	ɸnt	ɸmɸ	ɸaʊ	ɸmɸ	əʊə	ɸnd	ɸrɸ
	ɸaʊɸ	wɸ	æɸ	ɸr	ɸp	ɸns	klɸ	ɸŋ	klɸ
	ɸst	ɸf	ɸdɸ	ɸiɸ	ɸʌɸ	ɸkɸ	plɸ	ɸftɸ	ɸɸɸ
	ɸɔ:	ɸs	ɸts	wɸ	ɸu:ɸ	ɸɔiɸ	ɸmbɸ	ɸaʊəɸ	ɸiəɸ
	tɸ	tɸ	ɸð	dʒɸ	nɸ	ɸfʃɸ	eə	ɸtl	ɸndʒ

Topics	Topic 8	Topic 11	Topic 16	Topic 0	Topic 9	Topic 7	Topic 1	Topic 6
Alpha	0.80654	0.75588	0.63136	0.47514	0.45539	0.27874	0.18092	0.09079
Key (1-20)	hɸ	ai	ɸʌɸ	ðɸ	jɸ	fɸ	ɸeɸ	lɸ
	ɸz:	mɸ	sɸ	ɸə	ɸəɸ	ɸa:	ɸiɸ	ɸi:
	ɸd	tɸ	ɸeɸ	jɸ	ɸɔ:	ɸθ	ɸɔ:ɸ	ɸeəɸ
	ʃɸ	ə	ɸðɸ	ɸz	ɸr	əɸ	ɸm	ɸk
	ɸz:ɸ	ɸu:	ɸgɸ	əʊ	əʊɸ	a:	fɸ	ɸr
	ɸiɸ	ɸai	ɸəɸ	ɸeəɸ	ɸts	i:ɸ	rɸ	ɸrɸ
	lɸ	eɸ	ɸŋ	ɸaiɸ	ɸdɸ	ɸɔi	ɸth	ɸŋglɸ
	ɸa:ɸ	ɸeəɸ	əɸ	ɸvɸ	aiɸ	glɸ	ɸi	a:
	ɸeiɸ	əʊ	ʌɸ	ɸeɸ	ɸi:ɸ	ɸei	bɸ	ɸŋfʃ
	ɸps	ɸft	lɸ	ɸŋk	ɸnz	ɸaiəɸ	ɸmɸ	ɸiaɸ
	ɸθ	ɸwɸ	ɸu:	tɸ	ɸdʒɸ	ɸsɸ	ɸaɸ	frɸ
	ɸð	ɸvd	ɸr	ɸaʊə	ɸndʒɸ	ɸiəɸ	ɸflɸ	ɸlnɸ
	ɸbɸ	ɸsɸ	ɸnlɸ	ɸʊɸ	eɪɸ	ɸmbɸ	stɸ	eɪ
	ɸlf	ɸnd	ɸiə	ɸuɸ	ɸuəɸ	ɸvnɸ	ɸdɸ	ɸmpl
	ɸld	lɸ	ŋjɸ	smɸ	ɸðd	welaway	gɸ	ɸʊ
	grɸ	kɸ	ɸkt	ɸe	ɸæɸ	ɸdnz	ɸnz	ɸgmɸ
	aiɸ	ɸnsɸ	a:	swɸ	ɸtl	əʊɸ	rosalind	ɸvl
	ɸeɸ	ɸvlɸ	ɸdn	ɸæ	əʊəɸ	ə	ɸspɸ	ɸʃ
	drɸ	aiɸ	ɸmɸ	ɸskɸ	ɸspɸ	frɸ	ɸshɸ	ɸaiɸ
	ɸnfɸ	ɸk	skɸ	ɸltɸ	catullus	ɸʊ	ɸtl	prythee

それぞれのトピックがどのような「意味」で分類されているのかを検討するために、各トピックに母音、子音の語頭音 (Xɸ), 語中音 (ɸXɸ), 語末音 (ɸX) がそれぞれ何音含まれているかを数え

たものを表 8 に示す。なお、語中音のみ、一語に複数音現れる可能性がある。表 8 をみると、いくつかのトピックの特徴が見て取れる。Topic 3, 10, 4 のキーワードには、語中子音 (C-Middle) が現れておらず、子音は語頭音 (C-Head) と語末音 (C-Tail) が多く現れており、語中子音が 0 である代わりに語中母音 (V-Middle) が複数現れている。このことから、この三つのトピックは CVC (子音-母音-子音), VC, CV で構成される語の要素を多く含むトピックとみることができる。

表 7 raw テキストの単語頻度上位 20 位と単語を構成する音

Rank	Word	Sound	Word FREQ	Rank	Word	Sound	Word FREQ
1	the	<u>ð</u> θ φə (941) / <u>ð</u> θ φi (52)	993	12	s	φz (877) / φs (372)	127
2	and	<u>æ</u> θ <u>ɒ</u> <u>n</u> d	615	13	is	ɪθ φz	120
3	of	ɒθ φv	451	14	for	<u>f</u> θ φɔ: (79) / <u>f</u> θ φɔ:θ φr (31)	110
4	to	<u>t</u> θ φu: (289)	256	15	my	mθ φaɪ	109
5	in	ɪθ <u>ɒ</u> n	247	16	all	ɔ:θ <u>ɔ</u> l	103
6	a	ə (229) / eɪ (4)	233	17	thy	<u>ð</u> θ φaɪ	101
7	that	<u>ð</u> θ φæθ <u>θ</u> t	196	18	but	bθ φ <u>ʌ</u> θ <u>θ</u> t	96
8	i	aɪ	174	19	not	nθ φɒθ <u>θ</u> t	94
9	with	<u>w</u> θ φɪθ φð	165	20	from	<u>f</u> θ φrθ φɒθ <u>ɒ</u> m	91
10	d	θd (632) / <u>θ</u> t (1098)	145	20	so	sθ φəʊ	91
11	her	<u>h</u> θ φɜ: (115) / <u>h</u> θ φɜ:θ φr (19)	134				

* 下線は Topic 3 のキーワードに現れている音

** Sound 欄の () 内の数字は IPA sep テキストにおいて対応する音の頻度

表 8 IPA sep テキストの LDA によるトピックごとの音数比較 (α値降順)

Topic	Alpha	V-Head	V-Middle	V-Tail	C-Head	C-Middle	C-Tail
3	6.10462	1	6	3	4	0	6
10	3.61320	2	3	2	8	0	5
14	3.44713	2	3	1	5	1	8
12	2.93176	1	3	1	8	2	5
4	2.54886	0	4	3	7	0	6
13	1.86466	1	5	1	3	7	3
2	1.72640	0	6	4	4	1	3
5	1.42596	3	4	1	3	6	3
15	1.09778	1	5	1	6	3	4
8	0.80654	1	5	1	5	2	6
11	0.75588	2	1	2	3	5	4
16	0.63136	2	3	2	4	4	4
0	0.47514	0	5	4	5	3	2
9	0.45539	4	4	1	1	4	5
7	0.27874	3	2	4	3	3	2
1	0.18092	0	4	1	5	5	3
6	0.09079	0	3	2	2	4	6

* V : 母音, C : 子音, -Head : 語頭音, -Middle : 語中音, -Tail : 語末音



図3 ヒートマップ (IPA sep テキスト)

Topic 0,1,6 は語頭母音 (V-Head) が 0 であるものの、それ以外の音は全て現れているため、CVC(V) のように語頭が子音で始まり、後続する母音あるいは子音が繰り返される語を構成する音を多く含んでいるトピックと言える。また、17 トピックのうち 15 トピックにおいて、語中の母音は 3 つ以上出現しているが、Topic 11 と Topic 7 にはそれぞれ φεαφ と φαιαφ, φιαφ いずれ

も二重母音しか現れておらず，これらの二重母音と子音との間の関係によってトピックが分類されているとも考えられる。

以下に示す図3は，IPA sep テキストの LDA 実行結果のヒートマップである。これまでのヒートマップと同様，より黒いセルは，そのトピックが作品に顕著に現れることを示している。図3の上部にあるデンドログラムは，各トピックを分類している。このデンドログラムに従い，並べ替えた各トピックに語のどの位置にある音がいくつ含まれているかを数えたものが表9である。

図3及び表9のトピックを分類するデンドログラムでは，17トピックは表9上部の Topic 0 から Topic 8 までと，下部の Topic 3 から Topic 16 までに大きく分けられる。下部に分類される9つのトピックのうち， α 値 ≥ 1.0 は6トピックであるのに対し，上部の8トピック中， α 値 ≥ 1.0 は3トピックであり，下部のクラスターがより α 値が高く，複数の作品に顕著に現れるトピックを有していると言える。この表9より，上部，下部の各クラスター中，語頭音・語中音・語末音のそれぞれが最も多くキーワードに出現しているトピック（網掛け部）は6つあるが，以下で上部の3トピック (Topic 4, 1, 12) を raw テキストと比較しながら考察する。

表 9 IPA sep テキストのトピック毎の音数比較（デンドログラム順）

Topic	Alpha	V-Head	V-Middle	V-Tail	C-Head	C-Middle	C-Tail	Head	Middle	Tail
0	0.47514	0	5	4	5	3	2	5	8	6
6	0.09079	0	3	2	2	4	6	2	7	8
4	2.54886	0	4	3	7	0	6	7	4	9
15	1.09778	1	5	1	6	3	4	7	8	5
1	0.18092	0	4	1	5	5	3	5	9	4
12	2.93176	1	3	1	8	2	5	9	5	6
7	0.27874	3	2	4	3	3	2	6	5	6
8	0.80654	1	5	1	5	2	6	6	7	7
3	6.10462	1	6	3	4	0	6	5	6	9
9	0.45539	4	4	1	1	4	5	5	8	6
10	3.61320	2	3	2	8	0	5	10	3	7
5	1.42596	3	4	1	3	6	3	6	10	4
13	1.86466	1	5	1	3	7	3	4	12	4
2	1.72640	0	6	4	4	1	3	4	7	7
14	3.44713	2	3	1	5	1	8	7	4	9
11	0.75588	2	1	2	3	5	4	5	6	6
16	0.63136	2	3	2	4	4	4	6	7	6

Topic 12 が最も顕著に現れる作品は，Sonnt-B132 (“Sonnet [Blow ye the trumpet gather from afar]”, 1832) である。以下 (4) には IPA sep テキストの Sonnt-B132 に，Topic 12 のキーワードを網掛けしたもの，(5) には raw テキストを左側に，IPA テキストを右側に示し，中央には行番号を付したものを示す。まず，語頭の /ð/ 音は，定冠詞 the や指示代名詞 that, this など用途が広く，他の複数のトピックのキーワードにもなっているが，Sonnt-B132では，先にあげた the, this のほか，those, そして than が3回使われている。この3回のうち2回は行頭で ‘Than when’ と，when とともに使われており，語末音 /n/ もまた，この Topic 12 のキーワードである。(5) の10, 12, 14行目では，行末でこの語末音 /n/ で押韻している。他にキーワードに含まれている脚韻の要素には /əʊ/ と /ɪd/ があり，2-3行目，6-7行目で使用されている。語中音 /əʊ/ は3行目の中だけで3回使われている。3行目冒頭2語の二重母音 /aɪ/ と /eɪ/ は Topic 12 のキーワードにはなっていないが，

非円唇低母音－非円唇高母音 /aɪ/ から非円唇中母音－非円唇高母音 /eɪ/ そして非円唇と円唇の中間で発音される弱母音－円唇高母音 /əʊ/ へと、二重母音をたたみかけると同時に非円唇を円唇に、低母音から高母音へと調音位置を移動させることにより、音を後続の /əʊ/ と呼応させるだけでなく、躍動感を持たせる効果があると考えられる。

(4)

blɒ ɸəʊ jɸ ɸi: ɸɔ ɸə trɒ ɸʌɸ ɸmpɸ ɸɪɸ ɸt ɡɸ ɸæɸ ɸðɸ ɸə frɒ ɸvɸ ɸm əɸ ɸfɸ ɸɑ:
 ɸɔ ɸə hɸ ɸəʊɸ ɸsts tɸ ɸu: bɸ ɸæɸ ɸtl bɸ ɸi: nɸ ɸvɸ ɸt bɸ ɸɑ:ɸ ɸt æɸ ɸnd sɸ ɸəʊɸ ɸld
 əɸ ɸrɸ ɸaɪɸ ɸz brɸ ɸeɪɸ ɸv ɸɸ ɸəʊɸ ɸlz ɸɔ ɸə bɸ ɸəʊɸ ɸldɸ ɸɪɸ ɸst ɸɸ ɸv ɸɔ ɸə bɸ ɸəʊɸ ɸld
 brɸ ɸeɪɸ ɸk θrɸ ɸu: jɸ ɸɑ:ɸ ɸr aɪəɸ ɸn ʃɸ ɸæɸ ɸklz flɸ ɸɪɸ ɸŋ ɸɔ ɸeɸ ɸm fɸ ɸɑ:
 əʊ fɸ ɸɑ: ɸɔ ɸəʊɸ ɸz dɸ ɸeɪɸ ɸz ɸɸ ɸv Pɸ ɸiaɸ ɸst eə ɸɔ ɸə zɸ ɸɑ:
 ɡrɸ ɸu: tɸ ɸu: ɸɔ ɸɪɸ ɸs strɸ ɸeɸ ɸŋθ əɸ ɸmɸ ɸʌɸ ɸŋ hɸ ɸɪɸ ɸz dɸ ɸeɸ ɸzɸ ɸæɸ ɸts kɸ ɸəʊɸ ɸld
 wɸ ɸeɸ ɸn i:ɸ ɸvɸ ɸæɸ ɸn tɸ ɸu: mɸ ɸvɸ ɸskɸ ɸəʊɸ ɸz kjɸ ɸu:ɸ ɸpɸ ɸæɸ ɸlɸ ɸæɸ ɸz wɸ ɸɑ: rɸ ɸəʊɸ ɸld
 ɸɔ ɸə ɡrɸ ɸəʊɸ ɸŋ mɸ ɸɑ:ɸ ɸmɸ ɸæɸ ɸz ɸɸ ɸv ɸɔ ɸə pɸ ɸvɸ ɸlɸ ɸɪɸ ɸf wɸ ɸɑ:
 nɸ ɸəʊ mɸ ɸʌɸ ɸst jɸ ɸɑ: nɸ ɸəʊɸ ɸbl æɸ ɸŋɡɸ ɸə blɸ ɸeɪɸ ɸz əʊɸ ɸt mɸ ɸɑ:
 ɸɔ ɸæɸ ɸn wɸ ɸeɸ ɸn frɸ ɸvɸ ɸm Sobieski klɸ ɸæɸ ɸn bɸ ɸaɪ klɸ ɸæɸ ɸn
 ɸɔ ɸə mɸ ɸvɸ ɸzlɸ ɸeɸ ɸm mɸ ɸɪɸ ɸrɸ ɸɪəɸ ɸdz fɸ ɸeɸ ɸl æɸ ɸnd flɸ ɸeɸ ɸd bɸ ɸɪɸ ɸfɸ ɸɑ:
 ɸɔ ɸæɸ ɸn wɸ ɸeɸ ɸn Zamoysky smɸ ɸəʊɸ ɸt ɸɔ ɸə tɸ ɸɑ:ɸ ɸtɸ ɸə kɸ ɸɑ:ɸ ɸn
 ɸɔ ɸæɸ ɸn z:ɸ ɸlɸ ɸɪə wɸ ɸeɸ ɸn ɸɸ ɸn ɸɔ ɸə bɸ ɸɑ:ɸ ɸltɸ ɸɪɸ ɸk ʃɸ ɸɑ:
 Boleslas drɸ ɸəʊɸ ɸv ɸɔ ɸə pɸ ɸvɸ ɸmɸ ɸæɸ ɸrɸ ɸeɪɸ ɸŋjɸ ɸæɸ ɸn

(5)

Blow ye the trumpet, gather from <u>afar</u>	1	bləʊ jɪ: ɸə 'trʌmpɪt, 'gæðə frɒm ə 'fɑ:
The hosts to battle: be not bought and <u>sold</u> .	2	ðə həʊsts tu: 'bætl: bi: nɒt bɔ:t ənd səʊld.
<i>Arise, brave Poles, the boldest of the bold;</i>	3	ə 'raɪz, breɪv pəʊlz, ðə 'bəʊldɪst ɒv ðə bəʊld;
Break through your iron shackles fling them <u>far</u> .	4	breɪk θru: jɔ:r 'aɪən 'ʃækklz flɪŋ ðem fɑ:
O for those days of Piast, ere the <u>Czar</u>	5	əʊ fɔ: ɸəʊz deɪz ɒv Piast, eə ðə zɑ:
Grew to this strength among his deserts <u>cold</u> ;	6	ɡru: tu: ðɪs strɛŋθ ə 'mʌŋ hɪz 'deɪzəts kəʊld;
When even to Moscow's cupolas were <u>rolled</u>	7	wen 'i:vən tu: 'mɒskəʊz 'kju:pələz wɜ: rəʊld
The growing murmurs of the Polish war!	8	ðə 'grəʊŋ 'mɜ:məz ɒv ðə 'pɒlɪʃ wɔ:
Now must your noble anger blaze out <u>more</u>	9	nəʊ mʌst jɔ: 'nəʊbl 'æŋgə bleɪz əʊt mɔ:
<i>Than when from Sobieski, clan by <u>clan</u>,</i>	10	ðæn wen frɒm Sobieski, klæn baɪ klæn,
The Moslem myriads fell, and fled <u>before</u>	11	ðə 'mɒzləm 'mɪrɪədz fɛl, ənd fled brɪ 'fɔ:
<i>Than when Zamoysky smote the Tartar <u>Khan</u>,</i>	12	ðæn wen Zamoysky sməʊt ðə 'tɑ:tə kɑ:n,
<i>Than earlier, when on the Baltic <u>shore</u></i>	13	ðæn 'ɜ:lə, wen ɒn ðə 'bɔ:lɪk ʃɔ:
Boleslas drove the <u>Pomeranian</u> .	14	Boleslas drəʊv ðə 'pɒmə'reɪnjən.

“Sonnet [Blow ye the trumpet gather from afar]” (1832)

上記 (5) の中で押韻が確認できるものの、Topic 12 のキーワードに含まれていない要素に、1, 4, 5行目の行末の母音 /ɑ:/ (ɸɑ:) と、8, 9, 11, 13行目の行末音 /ɔ:/ (ɸɔ:) がある。ɸɑ: は Topic 7 ($\alpha = 0.27874$) のキーワードであり、ɸɔ: は Topic 3 ($\alpha = 6.10462$) と Topic 9 ($\alpha = 0.45539$) のキーワードである。Topic 7 が顕著に現れる作品の上位5作品は、上位から順に Farfrwy89 (“Far far

away”, 1889), Songvry30 (“Song [Every day hath its night]”, 1830), ThSlpngBe42 (“The Sleeping Beauty”, 1842), ThSlpng30 (“The Sleeping Beauty”, 1830), SCTCHSN27 (“Scotch song”, 1827) だが, 語末の /ɑ:/ (φɑ:) が最も多く用いられているのは, Farfrwy89 であり, 以下 (6) に引用する raw/IPA テキストからも分かるように, この φɑ: は押韻することを主目的として使われたというよりは, 寧ろ far という語を反復 (repetition) させることによって得られる効果 (スウィングしているようなリズムを生み出すなど) を狙ったことにより同じ語が繰り返し使われ, 結果として同じ音が多用されたと考えられる。この作品内で far という語が複数回用いられているだけではなく, φɑ: の前にくる音は語頭音と語中音合わせて7種類45例であり, そのうち語頭音 fφ は28例, 次に多いのは stφ の7例と大きく差が開いており, 必然的に φɑ: とともに far を構成する fφ も, このトピックのキーワードである。

(6)

WHAT SIGHT so lured him thro' the fields he knew	1	wɒt saɪt səʊ lʃʊəd hɪm θru: ðə fi:ldz hi: nju:
As where earth's green stole into heaven's own hue,	2	æz weə ɜ:θs gri:n stəʊl 'ɪntu: 'hevnz əʊn hju:,
Far — far — away?	3	fɑ: — fɑ: — ə'weɪ?
What sound was dearest in his native dells?	4	wɒt saʊnd wɒz 'diərist ɪn hɪz 'neɪtɪv delz?
The mellow lin-lan-lone of evening bells	5	ðə 'meləʊ lɪn-læn-ləʊn ɒv 'i:vnɪŋ belz
Far — far — away.	6	fɑ: — fɑ: — ə'weɪ.
What vague world-whisper, mystic pain or joy,	7	wɒt veɪg wɜ:ld-'wɪspə, 'mɪstɪk peɪn ɔ: dʒɔɪ,
Thro' those three words would haunt him when a boy,	8	θru: ðəʊz θri: wɜ:dz wʊd haʊnt hɪm wen ə bɔɪ,
Far — far — away?	9	fɑ: — fɑ: — ə'weɪ?
A whisper from his dawn of life? a breath	10	ə 'wɪspə frɒm hɪz dɔ:n ɒv laɪf? ə breθ
From some fair dawn beyond the doors of death	11	frɒm sʌm feə dɔ:n brɪ'jɒnd ðə dɔ:z ɒv deθ
Far — far — away?	12	fɑ: — fɑ: — ə'weɪ?
Far, far, how far? from o'er the gates of Birth,	13	fɑ:, fɑ:, haʊ fɑ: frɒm 'əʊə ðə geɪts ɒv bɜ:θ,
The faint horizons, all the bounds of earth,	14	ðə feɪnt hɔ:'raɪznz, ɔ:l ðə baʊndz ɒv ɜ:θ,
Far — far — away?	15	fɑ: — fɑ: — ə'weɪ?
What charm in words, a charm no words could give?	16	wɒt tʃɑ:m ɪn wɜ:dz, ə tʃɑ:m nəʊ wɜ:dz kʊd gɪv?
O dying words, can Music make you live	17	əʊ 'daɪɪŋ wɜ:dz, kæn 'mjuzɪk meɪk ju: lɪv
Far — far — away?	18	fɑ: — fɑ: — ə'weɪ?

(“Far far away”, 1889)

Topic 3 ($\alpha = 6.10462$) と Topic 9 ($\alpha = 0.45539$) のキーワードである φɑ: は, それぞれのトピックが顕著に現れる上位5作品を見ても, 語末音の φɑ: が脚韻として用いられている例は見つからなかった。語末音 φɑ: が二つのトピックに共通して現れる位置の前後には, /i:/ や /u:/ などの長母音や, 二重母音が現れていることが多いようにも見えるが, 音韻的な理由で語末音の φɑ: が

選択されたのか、または語末音 $\varphi\alpha:$ を含む語に高頻度語が多いためにキーワードとなったのかを判断するには、更に詳細な考察が必要である。

Topic 1 が最も顕著に現れるのは、(3) にも引用した、RiflmmF92 (“Riflemen Form”, 1892) である。RiflmmF92 の1-12行までに、Topic 1のキーワード20位までの語を色付けしたものを以下 (8) に示す。Raw テキストの LDA の結果で同じトピックに分類された、storm と form を構成する音 (st φ , f φ , $\varphi\alpha:$ φ , $\varphi\alpha\text{m}$) は、本 IPA sep テキストの結果においても全てキーワードに入っている。

(7)

$\delta\varphi$ $\varphi\alpha\varphi$ φr $\text{i}\varphi$ φz ə $\text{s}\varphi$ $\varphi\alpha\text{u}\varphi$ φnd $\text{v}\varphi$ φv $\theta\varphi$ $\varphi\text{L}\varphi$ $\varphi\text{nd}\varphi$ $\varphi\alpha\varphi$ φr $\text{ə}\varphi$ $\varphi\text{f}\varphi$ $\varphi\alpha:$
st φ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$ $\text{i}\varphi$ φn $\delta\varphi$ $\varphi\alpha$ $\text{s}\varphi$ $\varphi\alpha\text{u}\varphi$ $\varphi\theta$ $\delta\varphi$ $\varphi\alpha\varphi$ φt $\text{d}\varphi$ $\varphi\alpha:$ φ $\varphi\text{k}\varphi$ $\varphi\alpha\varphi$ φnz $\delta\varphi$ $\varphi\alpha$ $\text{d}\varphi$ $\varphi\text{e}\text{i}$
st φ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$ $\text{v}\varphi$ φv $\text{b}\varphi$ $\varphi\alpha\varphi$ φfl $\text{æ}\varphi$ φnd $\theta\varphi$ $\varphi\text{L}\varphi$ $\varphi\text{nd}\varphi$ $\varphi\alpha\varphi$ φr $\text{v}\varphi$ φv $\text{w}\varphi$ $\varphi\alpha:$
 $\text{w}\varphi$ $\varphi\text{e}\varphi$ φl $\text{i}\varphi$ φf $\text{i}\varphi$ φt $\text{d}\varphi$ $\varphi\text{u}:$ $\text{n}\varphi$ $\varphi\text{v}\varphi$ φt $\text{r}\varphi$ $\varphi\alpha\text{u}\varphi$ φl $\text{a}\text{u}\text{ə}$ $\text{w}\varphi$ $\varphi\text{e}\text{i}$
st φ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$ st φ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$ $\text{R}\varphi$ $\varphi\text{i}\varphi$ $\varphi\text{fl}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{m}\varphi$ $\varphi\text{e}\varphi$ φn $\text{f}\varphi$ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$
 $\text{r}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{d}\varphi$ φi $\text{b}\varphi$ $\varphi\text{i}:$ $\text{r}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{d}\varphi$ φi $\text{ə}\varphi$ $\varphi\text{g}\varphi$ $\varphi\text{e}\varphi$ φnst $\delta\varphi$ $\varphi\alpha$ st φ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$
 $\text{R}\varphi$ $\varphi\text{i}\varphi$ $\varphi\text{fl}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{m}\varphi$ $\varphi\text{e}\varphi$ φn $\text{R}\varphi$ $\varphi\text{i}\varphi$ $\varphi\text{fl}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{m}\varphi$ $\varphi\text{e}\varphi$ φn $\text{R}\varphi$ $\varphi\text{i}\varphi$ $\varphi\text{fl}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{m}\varphi$ $\varphi\text{e}\varphi$ φn $\text{f}\varphi$ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$
 $\text{b}\varphi$ $\varphi\text{i}:$ $\text{n}\varphi$ $\varphi\text{v}\varphi$ φt $\text{d}\varphi$ $\varphi\text{e}\varphi$ φf $\text{t}\varphi$ $\varphi\text{u}:$ $\delta\varphi$ $\varphi\alpha$ $\text{s}\varphi$ $\varphi\alpha\text{u}\varphi$ φnd $\delta\varphi$ $\varphi\alpha\varphi$ φt $\text{w}\varphi$ $\varphi\alpha:$ φ φnz
 $\text{b}\varphi$ $\varphi\text{i}:$ $\text{n}\varphi$ $\varphi\text{v}\varphi$ φt $\text{g}\varphi$ $\varphi\text{L}\varphi$ φld $\text{b}\varphi$ $\varphi\alpha\text{i}$ ə $\text{d}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{sp}\varphi$ $\varphi\text{v}\varphi$ φts $\text{pl}\varphi$ $\varphi\text{i}:$
 $\alpha:$ $\text{f}\varphi$ $\varphi\text{i}\varphi$ $\varphi\text{g}\text{z}$ $\text{v}\varphi$ φv $\theta\varphi$ $\varphi\text{r}\varphi$ $\varphi\text{sl}\text{z}$ $\text{ɔ}:$ $\text{gr}\varphi$ $\varphi\text{e}\text{i}\varphi$ φps $\text{v}\varphi$ φv $\theta\varphi$ $\varphi\alpha:$ φ φnz
 $\text{h}\varphi$ $\varphi\alpha\text{u}$ $\text{k}\varphi$ $\varphi\alpha\varphi$ φn ə $\text{d}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{sp}\varphi$ $\varphi\text{v}\varphi$ φt $\text{f}\varphi$ $\varphi\text{i}:\varphi$ φl $\text{w}\varphi$ $\varphi\text{r}\varphi$ $\varphi\delta$ $\delta\varphi$ $\varphi\alpha$ $\text{fr}\varphi$ $\varphi\text{i}:$
 $\text{f}\varphi$ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$ $\text{f}\varphi$ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$ $\text{R}\varphi$ $\varphi\text{i}\varphi$ $\varphi\text{fl}\varphi$ $\varphi\text{e}\varphi$ $\varphi\text{m}\varphi$ $\varphi\text{e}\varphi$ φn $\text{f}\varphi$ $\varphi\alpha:$ φ $\varphi\alpha\text{m}$

(8)

THERE is a sound of thunder afar,	1	$\delta\text{e}\text{ər}$ $\text{ɪ}\text{z}$ ə $\text{s}\alpha\text{ʊ}\text{nd}$ vv $'\theta\text{ʌ}\text{nd}\text{ə}\text{r}$ $\text{ə}'\text{f}\alpha:$,
Storm in the South that darkens the day!	2	$\text{s}\text{t}\alpha:\text{m}$ $\text{ɪ}\text{n}$ $\delta\text{ə}$ $\text{s}\alpha\text{ʊ}\theta$ $\delta\text{æ}\text{t}$ $'\text{d}\alpha:\text{k}\text{ə}\text{n}\text{z}$ $\delta\text{ə}$ $\text{d}\text{e}\text{i}!$
Storm of battle and thunder of war!	3	$\text{s}\text{t}\alpha:\text{m}$ vv $'\text{b}\text{æ}\text{t}\text{l}$ $\text{ə}\text{nd}$ $'\theta\text{ʌ}\text{nd}\text{ə}\text{r}$ vv $\text{w}\alpha:$!
Well if it do not roll our way.	4	wel $\text{ɪ}\text{f}$ $\text{ɪ}\text{t}$ $\text{d}\text{u}:$ $\text{n}\text{ɒ}\text{t}$ $\text{r}\alpha\text{ʊ}\text{l}$ $'\text{a}\text{ʊ}\text{ə}$ wei .
Storm, Storm, Riflemen form!	5	$\text{s}\text{t}\alpha:\text{m}$, $\text{s}\text{t}\alpha:\text{m}$, $\text{R}\text{ɪ}\text{f}\text{l}\text{e}\text{m}\text{e}\text{n}$ $\text{f}\alpha:\text{m}!$
Ready, be ready against the storm!	6	$'\text{r}\text{e}\text{d}\text{i}$, $\text{b}\text{i}:$ $'\text{r}\text{e}\text{d}\text{i}$ $\text{ə}'\text{g}\text{e}\text{n}\text{s}\text{t}$ $\delta\text{ə}$ $\text{s}\text{t}\alpha:\text{m}!$
Riflemen, Riflemen, Riflemen form!	7	$\text{R}\text{ɪ}\text{f}\text{l}\text{e}\text{m}\text{e}\text{n}$, $\text{R}\text{ɪ}\text{f}\text{l}\text{e}\text{m}\text{e}\text{n}$, $\text{R}\text{ɪ}\text{f}\text{l}\text{e}\text{m}\text{e}\text{n}$ $\text{f}\alpha:\text{m}!$
Be not deaf to the sound that warns,	8	$\text{b}\text{i}:$ $\text{n}\text{ɒ}\text{t}$ def $\text{t}\text{u}:$ $\delta\text{ə}$ $\text{s}\alpha\text{ʊ}\text{nd}$ $\delta\text{æ}\text{t}$ $\text{w}\alpha:\text{n}\text{z}$,
Be not gull'd by a despot's plea!	9	$\text{b}\text{i}:$ $\text{n}\text{ɒ}\text{t}$ $\text{g}\text{ʌ}\text{ld}$ $\text{b}\alpha\text{i}$ $\text{ə}'\text{d}\text{e}\text{s}\text{p}\text{ɒ}\text{t}\text{s}$ $\text{p}\text{l}\text{i}:$!
Are figs of thistles? or grapes of thorns?	10	$\alpha:$ $\text{f}\text{ɪ}\text{g}\text{z}$ vv $'\theta\text{ɪ}\text{s}\text{l}\text{z}?$ $\text{ɔ}:$ $\text{g}\text{r}\text{e}\text{ɪ}\text{p}\text{s}$ vv $\theta\alpha:\text{n}\text{z}?$
How can a despot feel with the Free?	11	$\text{h}\alpha\text{ʊ}$ $\text{k}\text{æ}\text{n}$ $\text{ə}'\text{d}\text{e}\text{s}\text{p}\text{ɒ}\text{t}$ $\text{f}\text{i}:\text{l}$ $\text{w}\text{ɪ}\delta$ $\delta\text{ə}$ $\text{f}\text{r}\text{i}:$?
Form, Form, Riflemen Form!	12	$\text{f}\alpha:\text{m}$, $\text{f}\alpha:\text{m}$, $\text{R}\text{ɪ}\text{f}\text{l}\text{e}\text{m}\text{e}\text{n}$ $\text{f}\alpha:\text{m}!$

“Riflemen Form” (1892: ll. 1–12)

(8) の raw/IPA テキストの8, 10行目には、同じ母音を使った脚韻があり (/wɔ:nz/, /θɔ:nz /), 一見すると押韻を検知したようにも見える。しかし、raw テキストで頻度を調べると、form はコーパス内で全28例中14例、storm は全12例中10例がこのRiflmmF92に生起しており、特にこの2語に共通する、 $\varphi\alpha:$ φ , $\varphi\alpha\text{m}$ は、この結果のみでは押韻を検知したと断言することはできない。更に、この2語に共通する $\varphi\alpha:$ φ を中心になる 3-gram の頻度 (頻度10以上) を以下表10に示す。

表 10 IPA sep テキストにおける $\varphi\alpha:\varphi$ を中心とする 3-gram (頻度 10 以上)

	Sound 1	NODE	Sound 2	FREQ
1	f φ	$\varphi\alpha:\varphi$	φr	31
2	f φ	$\varphi\alpha:\varphi$	φm	28
3	gl φ	$\varphi\alpha:\varphi$	$\varphi r\varphi$	23
4	m φ	$\varphi\alpha:\varphi$	φn	16
5	l φ	$\varphi\alpha:\varphi$	φd	15
6	$\theta\varphi$	$\varphi\alpha:\varphi$	φt	13
7	st φ	$\varphi\alpha:\varphi$	φm	12
8	w φ	$\varphi\alpha:\varphi$	$\varphi t\varphi$	12
9	d φ	$\varphi\alpha:\varphi$	φn	10
10	f φ	$\varphi\alpha:\varphi$	φl	10
11	f φ	$\varphi\alpha:\varphi$	$\varphi\theta$	10

Sound 1 の位置にくる音は、表10の11例中4例が $f\varphi$ であり、 $f\varphi$ - $\varphi\alpha:\varphi$ の組み合わせののべ頻度は 79/180である。また、頻度順2番目には $f\varphi\varphi\alpha:\varphi, \varphi m$ の28があり、 $\varphi\alpha:\varphi$ 音の前に位置し、語頭音として語を形成する音の中では $f\varphi$ が他の音よりも多く $\varphi\alpha:\varphi$ と結びつきやすいと言え、またこの語あるいは音同士の結びつきやすさや、組み合わせの頻度による影響を可能な限り避けた方が、押韻や音韻特徴に焦点を当てた分析ができると考える。この影響を避ける方法の一つとして、IPA 表記したテキストを更に分割する際に、本稿のように子音と母音で分けるのではなく、音節ごとに区切ることが挙げられるが、IPA テキストをどのように正確に音節ごとに分割するかなど課題も多く、稿を改めて議論したい。

最後に、Topic 4 を考察する。Topic 4 が最も顕著に現れる作品は、(9), (10) に挙げる MYBRTHR27 (“My brother”, 1827) である。この詩は四行連で書かれており、7つあるスタンザの

(9)

w φ $\varphi i\varphi$ φd $f\varphi$ $\varphi a\varphi$ $\varphi l r\varphi$ $\varphi i\varphi$ $\varphi n\eta$ $st\varphi$ $\varphi e\varphi$ φp $a i$ $k\varphi$ $\varphi e i\varphi$ φm $t\varphi$ $\varphi u:$ $s\varphi$ $\varphi i:$
 $i\varphi$ φn $d\varphi$ $\varphi e\varphi$ $\varphi\theta s$ $\lambda\varphi$ $\varphi n h\varphi$ $\varphi i:\varphi$ $\varphi d\varphi$ $\varphi i\varphi$ $\varphi\eta$ $\varepsilon\varphi$ $\varphi p\varphi$ $\varphi a\varphi$ $\varphi\theta\varphi$ φi
 $\delta\varphi$ $\varphi a\varphi$ φt $f r\varphi$ $\varphi e\varphi$ $\varphi n d$ $s\varphi$ $\varphi a\varphi$ $d\varphi$ $\varphi i a\varphi$ φr $i\varphi$ φn $l\varphi$ $\varphi a i\varphi$ φf $t\varphi$ $\varphi u:$ $m\varphi$ $\varphi i:$
m φ $\varphi a i$ $b r\varphi$ $\varphi l\varphi$ $\varphi d\varphi$ φa
 $\varphi i\varphi$ φd $f l\varphi$ $\varphi a u\varphi$ φz $v\varphi$ φv $l\varphi$ $\varphi l\varphi$ $\varphi v i\varphi$ $\varphi i i\varphi$ $\varphi s t$ $s\varphi$ $\varphi e\varphi$ $\varphi n t$ $\varepsilon\varphi$ $\varphi n d$ $h j\varphi$ $\varphi u:$
 $\delta\varphi$ $\varphi a\varphi$ φt $s t r\varphi$ $\varphi u:\varphi$ φd $\delta\varphi$ $\varphi a i$ $f\varphi$ $\varphi\alpha:\varphi$ φm $t v\varphi$ $\varphi v\varphi$ φz $s\varphi$ $\varphi a\varphi$ φd $t\varphi$ $\varphi u:$ $v j\varphi$ $\varphi u:$
 $\delta\varphi$ $\varphi a i$ $l\varphi$ $\varphi a i\varphi$ $\varphi p\varphi$ $\varphi i\varphi$ φs $f\varphi$ $\varphi e i\varphi$ φs $p\varphi$ $\varphi i:\varphi$ φp $w\varphi$ $\varphi v\varphi$ $\varphi n l\varphi$ φi $\theta r\varphi$ $\varphi u:$
m φ $\varphi a i$ $b r\varphi$ $\varphi l\varphi$ $\varphi d\varphi$ φa
w φ $\varphi a i$ $d\varphi$ $\varphi i\varphi$ φd $\delta\varphi$ $\varphi e i$ ($\delta\varphi$ $\varphi e\alpha$ $\delta\varphi$ $\varphi e i$ $d\varphi$ $\varphi i\varphi$ φd $n\varphi$ $\varphi v\varphi$ φt $f\varphi$ $\varphi i:\varphi$ φl)
w φ $\varphi i\varphi$ φd $s i j\varphi$ $\varphi u:\varphi$ $\varphi d\varphi$ $\varphi i a\varphi$ φs $k\varphi$ $\varphi e a\varphi$ φr $\alpha:\varphi$ φl $\varepsilon\varphi$ $\varphi l s$ $k\varphi$ $\varphi a\varphi$ $\varphi n s\varphi$ $\varphi i:\varphi$ φl
b φ $\varphi l\varphi$ φt $\delta\varphi$ $\varphi a i$ $k\varphi$ $\varphi a u\varphi$ $\varphi l d$ $f\varphi$ $\varphi e i\varphi$ φs $\varepsilon\varphi$ $\varphi l\varphi$ $\varphi a u\varphi$ φn $r\varphi$ $\varphi i\varphi$ $\varphi v\varphi$ $\varphi i:\varphi$ φl
m φ $\varphi a i$ $b r\varphi$ $\varphi l\varphi$ $\varphi d\varphi$ φa
 $\delta\varphi$ $\varphi e i$ $m\varphi$ $\varphi a i\varphi$ φt $h\varphi$ $\varphi a\varphi$ φv $n\varphi$ $\varphi a u\varphi$ φn $w\varphi$ $\varphi v\varphi$ φt $j\varphi$ $\varphi u:\varphi$ $\varphi z d$ $t\varphi$ $\varphi u:$ $g l\varphi$ $\varphi a\varphi$
w φ $\varphi i\varphi$ φd $s m\varphi$ $\varphi a i\varphi$ $\varphi l z$ $\varepsilon\varphi$ $\varphi n d$ $v\varphi$ $\varphi f t$ $d\varphi$ $\varphi i\varphi$ $\varphi s p\varphi$ $\varphi e\varphi$ $\varphi l d$ $m\varphi$ $\varphi a i$ $w\varphi$ $\varphi a\varphi$
w φ $\varphi u\varphi$ φd $f j\varphi$ $\varphi i\varphi$ φl $m\varphi$ $\varphi i:$ $m\varphi$ $\varphi a u\varphi$ $\varphi s t$ $w\varphi$ $\varphi e\varphi$ φn $f\varphi$ $\varphi e i\varphi$ $\varphi d\varphi$ $\varphi i\varphi$ φd $s\varphi$ $\varphi a\varphi$
m φ $\varphi a i$ $b r\varphi$ $\varphi l\varphi$ $\varphi d\varphi$ φa
 $\delta\varphi$ φa $t\varphi$ $\varphi a u\varphi$ $\varphi l\varphi$ $\varphi i\varphi$ $\varphi\eta$ $v\varphi$ φv $\delta\varphi$ $\varphi a i$ $f j\varphi$ $\varphi u:\varphi$ $\varphi n\varphi$ $\varphi a\varphi$ $\varphi r\varphi$ $\varphi a\varphi$ φl $b\varphi$ $\varphi e\varphi$ φl
 $\delta\varphi$ φa $n\varphi$ $\varphi a i\varphi$ φn $l\varphi$ $\varphi a\varphi$ $n\varphi$ $\varphi a u\varphi$ $\varphi t s$ $\delta\varphi$ $\varphi a\varphi$ φt $s\varphi$ $\varphi a u\varphi$ φk $\delta\varphi$ $\varphi a i$ $n\varphi$ $\varphi e\varphi$ φl
a i $n\varphi$ $\varphi a\varphi$ $n\varphi$ $\varphi v\varphi$ φt $h\varphi$ $\varphi a\varphi$ a i $b\varphi$ $\varphi\alpha:$ $s\varphi$ $\varphi a\varphi$ $w\varphi$ $\varphi e\varphi$ φl
m φ $\varphi a i$ $b r\varphi$ $\varphi l\varphi$ $\varphi d\varphi$ φa
b φ $\varphi l\varphi$ φt $\varepsilon\varphi$ $\delta\varphi$ φa $f j\varphi$ $\varphi i\varphi$ φl $d\varphi$ $\varphi a\varphi$ $\varphi\eta k$ $m\varphi$ $\varphi a u\varphi$ $\varphi l d$ $\delta\varphi$ $\varphi a\varphi$ φt $s l\varphi$ $\varphi i\varphi$ φd
 $\delta\varphi$ $\varphi l\varphi$ φl $s\varphi$ $\varphi a u\varphi$ $\varphi n d\varphi$ $\varphi i\varphi$ $\varphi\eta$ $v\varphi$ φn $\delta\varphi$ $\varphi a i$ $k\varphi$ $\varphi v\varphi$ $\varphi f\varphi$ $\varphi i\varphi$ φn $l\varphi$ $\varphi i\varphi$ φd
 $\delta\varphi$ $\varphi a\varphi$ φt $d r\varphi$ $\varphi u:$ $m\varphi$ $\varphi\alpha:$ $t\varphi$ $\varphi e a\varphi$ φz $\delta\varphi$ $\varphi a\varphi$ φn $\alpha:\varphi$ φl $b\varphi$ $\varphi i\varphi$ $\varphi s\varphi$ $\varphi a i\varphi$ φd
m φ $\varphi a i$ $b r\varphi$ $\varphi l\varphi$ $\varphi d\varphi$ φa
 $\varepsilon\varphi$ $\varphi n d$ $\delta\varphi$ $\varphi e\varphi$ φn a i $h\varphi$ $\varphi l\varphi$ $\varphi r\varphi$ $\varphi i\varphi$ φd $f\varphi$ $\varphi\alpha:\varphi$ $\varphi s t$ $\varepsilon\varphi$ $\varphi w\varphi$ $\varphi e i$
h φ $\varphi a\varphi$ $k\varphi$ $\varphi u\varphi$ φd a i $\varepsilon\alpha$ $h\varphi$ $\varphi a\varphi$ φv $b\varphi$ $\varphi\alpha:\varphi$ φn $t\varphi$ $\varphi u:$ $s t\varphi$ $\varphi e i$
w φ $\varphi e\alpha$ $k\varphi$ $\varphi e a\varphi$ $\varphi l\varphi$ $\varphi i\varphi$ φs $h\varphi$ $\varphi a\varphi$ $\varphi n d$ $i\varphi$ $\varphi n h j\varphi$ $\varphi u:\varphi$ $\varphi m d$ $\delta\varphi$ $\varphi a i$ $k l\varphi$ $\varphi e i$

(10)

WITH falt'ring step I came to <i>see</i> ,	1	wɪð falt'ring step aɪ kɛɪm tu: si:,
In Death's unheeding apathy,	2	ɪn deθs ʌn'hi:diŋ 'æpəθi,
That friend so dear in life to <i>me</i> ,	3	ðæt frend səʊ diər ɪn laɪf tu: mi:,
My brother!	4	maɪ 'brʌðə!
'Mid flowers of loveliest scent and hue	5	mɪd 'flaʊəz ɒv 'lʌvli:st sent ænd hju:
That strew'd thy form, 'twas sad to view	6	ðæt stru:d ðaɪ fɔ:m, twɒz sæd tu: vju:
Thy lifeless face peep wanly through,	7	ðaɪ 'laɪflɪs feɪs pi:p 'wɒnli θru:,
My brother!	8	maɪ 'brʌðə!
Why did they (there they did not <i>feel</i> !)	9	wai dɪd ðeɪ (ðeə ðeɪ dɪd nɒt fi:!)
With studious care all else <i>conceal</i> ,	10	wɪð 'stju:diəs keər ɔ:l els kən'si:l,
But thy cold face alone <i>reveal</i> ,	11	bʌt ðaɪ kəʊld feɪs ə'ləʊn rɪ'vi:l,
My brother!	12	maɪ 'brʌðə!
They might have known, what used to glow	13	ðeɪ maɪt hæv nəʊn, wɒt ju:zd tu: gləʊ
With smiles, and oft dispell'd my woe,	14	wɪð smaɪlz, ænd ɒft dɪs'peld maɪ wəʊ,
Would chill me most, when faded so,	15	wʊd tʃɪl mi: məʊst, wen 'feɪdɪd səʊ,
My brother!	16	maɪ 'brʌðə!
The tolling of thy funeral <i>bell</i> ,	17	ðə 'təʊlɪŋ ɒv ðaɪ 'fju:nərəl bɛl,
The nine low notes that spoke thy <i>knell</i> ,	18	ðə naɪnləʊ nəʊts ðæt spəʊk ðaɪ nɛl,
I know not how I bore so <i>well</i> ,	19	aɪ nəʊ nɒt haʊ aɪ bɔ: səʊ wɛl,
My brother!	20	maɪ 'brʌðə!
But oh! the chill, dank mould that <i>slid</i> ,	21	bʌt əʊ! ðə tʃɪl, dæŋk məʊld ðæt slɪd,
Dull-sounding, on thy coffin- <i>lid</i> ,	22	dʌl-'saʊndɪŋ, ɒn ðaɪ 'kɒfɪn-lɪd,
That drew more tears than all <i>beside</i> ,	23	ðæt dru: mɔ: teəz ðæn ɔ:l br'saɪd,
My brother!	24	maɪ 'brʌðə!
And then I hurried fast away;	25	ænd ðen aɪ 'hʌrɪd fɑ:st ə'weɪ;
How could I e'er have borne to stay	26	haʊ kʊd aɪ eə hæv bɔ:n tu: steɪ
Where careless hand inhumed thy clay,	27	weə 'keəlis hænd ɪn'hju:md ðaɪ kleɪ,
My brother!	28	maɪ 'brʌðə!

“My brother” (1827)

それぞれ1行目から3行目までが脚韻を踏んでいる。この7種の脚韻の音のうち、Topic 4 のキーワードに含まれているのは、部分的な (例えば、第3スタンザの脚韻は /i:l/ であるが、トピックのキーワードになっているのは子音の /l/ のみである) ものも含めて4種である。

この Topic 4 の α 値は 2.54886 であり、17あるトピックの中で5番目に高い数値であり、他の作品にも当該トピックのキーワードが現れている可能性が高い。そこで、MYBRTHR27 の次、つ

まり2番目と3番目に顕著に Topic 4 が現れる作品, THBRDSF27 (“The bards farewell”, 1827) (以下(11), (12)) と Sonnthb32 (“Sonnet [Oh beauty passing beauty]”, 1832) (以下(13), (14)) の2作品も考察する。

(11)

snφ φəuφ φdφ φəφ φn dφ φai klφ φiφ φfs fφ φəφ φl hφ φiə nφ φəu mφ φɔ:
dφ iφ φs dφ φi:φ φp tφ φəuφ φnd hφ φa:φ φp əφ φzφ φeφ φn
bφ φlφ φt bφ φəφ φnφ φə krφ φai əφ φnd bφ φəφ φtl rφ φɔ:
fφ φəφ φl fφ φɔ:φ φm ə fφ φiəφ φsφ φə strφ φeφ φn
əuə dφ φai swφ φi:φ φt kφ φɔ:φ φdz mφ φai mφ φəφ φɔ:φ φiφ φk lφ φaiə
wφ φvφ φt fjφ φu:φ φfφ φə hφ φəφ φnd fφ φəφ φl strφ φeφ
wφ φvφ φt brφ φeφ φn fφ φəφ φl fφ φi:φ φl dφ φai mφ φa:φ φstφ φəφ φz fφ φaiə
ɔ: frφ φeφ φm hφ φiφ φz mφ φəφ φfjφ φiφ φs lφ φeφ
wφ φeφ φl mφ φaiφ φt dφ φə krφ φa:φ φftφ φi εφ φdφ φəφ φd fφ φiə
fφ φuφ φd ai bφ φlφ φt tφ φlφ φfj dφ φai kφ φɔ:φ φd
iφ φts slφ φaiφ φtφ φiφ φst sφ φəuφ φnd wφ φuφ φd kφ φəuφ φfj dφ φə spφ φiə
əφ φnd bφ φeə dφ φi iφ iφ φndφ φiφ φzφ φəφ φnt sφ φɔ:φ φd
fφ φuφ φl wφ φeφ φl hφ φi: njφ φu: dφ φə wφ φiφ φzφ φəφ φd spφ φeφ φl
dφ φəφ φt dwφ φeφ φlt əφ φpφ φvφ φn dφ φai strφ φiφ φŋ
əφ φnd trφ φeφ φn hφ φi: hφ φɔ:φ φd dφ φai swφ φeφ φl
θrφ φu: snφ φəuφ φdφ φəφ φnz kφ φəφ φvφ φəφ φnz rφ φiφ φŋ
dφ φi:φ φz aiφ φz fφ φəφ φl slφ φi:φ φp iφ φn dφ φeφ φt φts dφ φlφ φl nφ φaiφ φt
dφ φiφ φs hφ φəφ φnd ɔ:φ φl nφ φɔ:φ φvlφ φiφ φs lφ φai
eə wφ φlφ φns əφ φzφ φeφ φn jφ φu: φn ɔ:φ φb vφ φv lφ φaiφ φt
brφ φeφ φk əuə dφ φə klφ φl φiə blφ φu: skφ φai
əφ φnd dφ φəu bφ φai hφ φeφ φlz əuφ φn fjφ φuəφ φrφ φiφ φz nφ φuφ φrst
lφ φnfφ φɔ:φ φl dφ φai bφ φəφ φnφ φəφ φz prφ φaiφ φd
bφ φlφ φt nφ φəu dφ φəφ φt lφ φiφ φvφ φiφ φŋ dφ φi: ai kφ φɔ:φ φst
əφ φnd kφ φɔ:φ φsφ φiφ φŋ dφ φi: ai dφ φaiφ φd

(12)

SNOWDON! thy cliffs shall hear no more	1	'snəʊdən! ðaɪ klɪfs ʃæl hiə nəʊ mɔː
This deep-toned harp again;	2	ðɪs di:p-təʊnd hɑ:p ə'gen;
But banner-cry and battle-roar	3	bʌt 'bænə-kraɪ ənd 'bætl-rɔː
Shall form a fiercer strain!	4	ʃæl fɔ:m ə 'fiəsə streɪn!
O'er thy sweet chords, my magic lyre!	5	'əʊə ðaɪ swi:t kɔ:dz, maɪ 'mædʒɪk 'laɪə!
What future hand shall stray?	6	wɒt 'fju:ʃə hænd ʃæl streɪ?
What brain shall feel thy master's fire,	7	wɒt breɪn ʃæl fi:l ðaɪ 'ma:stəz 'faɪə,
Or frame his matchless lay?	8	ɔ: freɪm hɪz 'mætʃlɪs leɪ?
Well might the crafty Edward fear:	9	wel maɪt ðə 'kra:ftɪ 'edwəd fiə:
Should I but touch thy chord,	10	ʃəd aɪ bʌt tʌʃ ðaɪ kɔ:d,
Its slightest sound would couch the spear,	11	ɪts 'slɑ:tɪst saʊnd wʊd kaʊʃ ðə spɪə,
And bare the indignant sword!	12	ænd beə ðɪ m 'dɪgnənt sɔ:d!
Full well he knew the wizard-spell	13	fʊl wel hi: nju: ðə 'wɪzəd-spel
That dwelt upon thy string;	14	ðæt dweɪlt ə'pɒn ðaɪ strɪŋ;
And trembled, when he heard thy swell	15	ænd 'tremblɪd, wen hi: hɜ:d ðaɪ swel
Thro' Snowdon's caverns ring!	16	θru: 'snəʊdənz 'kævənz rɪŋ!
These eyes shall sleep in death's dull night,	17	ði:z aɪz ʃæl sli:p ɪn deθs dʌl naɪt,
This hand all nerveless lie,	18	ðɪs hænd ɔ:l 'nɜ:vɪs laɪ,
Ere once again yon orb of light	19	eə wʌns ə'gen jɒn ɔ:b ɒv laɪt
Break o'er the clear blue sky!	20	breɪk 'əʊə ðə klɪə blu: skaɪ!

And thou, by Hell's own furies nurst,
 Unfurl thy banner's pride!
 But know that, living, thee I cursed;
 And, cursing thee, I died!

21 ænd ðau, bai helz əʊn 'fjuərəiz nurst,
 22 ,ʌn'fɜ:l ðai 'bænəz praɪd!
 23 bʌt nəʊ ðæt, 'lɪvɪŋ, ði: ai kɜ:st;
 24 ænd, 'kɜ:sɪŋ ði:, ai daɪd!

“The bards farewell” (1827)

(13)

au	bjɔ	əu:ə	ɔtə	ɔi	pə	əɑ:ə	ɔsə	ɔtə	ɔŋ	bjɔ	əu:ə	ɔtə	ɔi	swə	ɔi:ə	ɔtə	ɔtə	ɔst	swə	ɔi:ə	ɔt			
hə	əau	canst	ðə	əau	lɔ	ɔsə	ɔt	mə	ɔi:	wə	ɔsə	ɔst	mə	əai	ɔə	əu:ə	ɔt	lɔ	ɔn	sə	əaiə	ɔz		
ai	əuə	ɔnɔ	ɔi	ɑ:ə	ɔsk	tə	əu:	sə	ɔtə	ɔt	bə	ɔtə	ɔsə	əaiə	ɔd	ðə	əai	fə	ɔi:ə	ɔt				
ðə	əau	nə	ɔəuə	ɔst	ai	ðə	ɔsə	nə	ɔtə	ɔt	lɔ	ɔsə	ɔk	tə	ɔnɔ	əu:	ðə	əaiə	ɔn	aiə	ɔz			
mə	əai	ɔt	ai	bə	əauə	ɔt	kə	ɔtə	ɔs	ðə	əai	hə	əəə	ɔnd	ai	ðə	ɔsə	nə	ɔtə	ɔt	fə	ɔəuə	ɔld	
mə	əai	ɑ:ə	ɔmz	əp	ɔbə	əəuə	ɔt	ðə	ɔi:	skə	ɔsə	ɔst	ɔi	ðə	ɔsə	tə	əu:	spə	ɔi:ə	ɔk				
æp	ɔnd	nə	əauə	ɔbə	ɔtə	ɔŋ	sə	ɔi:ə	ɔmz	tə	əu:	mə	ɔi:	sə	ɔsə	wə	əaiə	ɔld	æp	ɔnd	bə	ɔəuə	ɔld	
æp	ɔz	wə	ɔtə	ɔd	wə	əauə	ɔn	kə	ɔtə	ɔs	tə	əu:	tə	əauə	ɔf	ðə	əai	blə	əəə	ɔst	fə	ɔi:ə	ɔk	
mə	ɔtə	ɔbə	ɔtə	ɔŋks	lɔ	ɔf	ai	fə	ɔsə	ɔd	kə	ɔtə	ɔs	ðə	ɔi:	nə	ɔsə	kə	ɔsə	ɔntɔ	ɔəuə	ɔl		
wə	ɔtə	ɔbə	ɔtə	ɔn	ðə	ɔs	θrə	ɔtə	ɔlə	ɔtə	ɔŋ	brə	ɔsə	ən	kə	ɔsə	ɔd	kə	ɔi:ə	ɔp	əp	əp	ɔəuə	ɔt
ðə	ɔs	sə	əauə	ɔtl	spə	ɔtə	ɔrə	ɔtə	ɔt	l:ə	ɔsə	ɔt	mə	ɔsə	ɔd	mə	əai	lɔ	əaiə	ɔs	sə	ɔəuə	ɔl	
ðə	ɔs	bə	ɔsə	wə	ɔz:ə	ɔd	kə	ɔtə	ɔs	hə	əəə	ɔt	mə	ɔsə	ɔd	mə	əai	lɔ	əaiə	ɔs	sə	ɔəuə	ɔl	
tə	əu:	tɔ	ɔsə	ɔmbl	lɔ	əaiə	ɔk	ə	lɔ	əu:ə	ɔtstɔ	ɔtə	ɔŋ	ɔs	ðə	ɔs	nə	ɔsə	ɔt					
hə	əəə	ɔt	mə	ɔsə	ɔtə	ɔtə	ɔd	lɔ	ən	ðə	ɔs	sə	əaiə	ɔlə	ɔsə	ɔns	ðə	əəə	ɔt	lɔ	ɔt	brə	ɔəuə	ɔk

(14)

Blow ye the trumpet, gather from afar
 The hosts to battle: be not bought and sold.
 Arise, brave Poles, the boldest of the bold;
 Break through your iron shackles fling them far.
 O for those days of Piast, ere the Czar
 Grew to this strength among his deserts cold;
 When even to Moscow's cupolas were rolled
 The growing murmurs of the Polish war!
 Now must your noble anger blaze out more
 Than when from Sobieski, clan by clan,
 The Moslem myriads fell, and fled before
 Than when Zamoysky smote the Tartar Khan,
 Than earlier, when on the Baltic shore
 Boleslas drove the Pomeranian.

1 bləʊ ji: ðə 'trʌmpɪt, 'gæðə frəm ə'fa:
 2 ðə hæʊsts tu: 'bætl: bi: nɒt bɔ:t ænd səʊld.
 3 ə'raɪz, breɪv pəʊlz, ðə 'bəʊldɪst ɒv ðə bəʊld;
 4 breɪk θru: jɔ:ɾ 'aɪən 'ʃæklz flɪŋ ðem fa:.
 5 əʊ fɔ: ðəʊz deɪz ɒv piast, eə ðə zɑ:
 6 grʊ: tu: ðɪs streɪŋθ ə'mʌŋ hɪz 'dezəts kəʊld;
 7 wen 'i:vən tu: 'mɒskəʊz 'kju:pələz wɜ: rəʊld
 8 ðə 'grəʊɪŋ 'mɜ:məz ɒv ðə 'pɒlɪʃ wɔ:!
 9 naʊ mʌst jɔ: 'nəʊbl 'æŋgə bleɪz aʊt mɔ:
 10 ðæn wen frəm sobieski, klæn baɪ klæn,
 11 ðə 'mɒzləm 'mɪrɪədz fɛl, ænd flɛd brɪ'fɔ:
 12 ðæn wen zamoyski sməʊt ðə 'tɑ:tə kɑ:n,
 13 ðæn 'z:lɪə, wen ɒn ðə 'bɔ:lɪk ʃɔ:
 14 boleslas drəʊv ðə ,pɒmə'reɪnjən.

“Sonnet [Oh beauty passing beauty]” (1832)

THBRDSF27 は、6スタンザからなる四行連の詩で abab のスタイル (1,3 行目と2,4 行目) で脚韻している。24行のうち、キーワードとしてマークされている部分は、先の MYBRTHR27 と比較すると少なく、3種であるが、そのうち、MYBRTHR27 と共通する語中音の母音 $\varphi\epsilon\varphi$ と、語末音の子音 φl (13, 15 行目), 及び語末音 φd (22, 24 行目) がみられる。もう一つの語末音に17, 19 行目の φt があるが、この脚韻は本来直前の母音 $\varphi aɪ\varphi$ も含めて然るべきであり、また22, 24 行目の脚韻も同様に、正確には直前の母音 $\varphi aɪ\varphi$ も語末音と合わせて扱うのが妥当であり、やは

り今後の研究では音節に分割して分析を進めることが課題であると言える。Sonnthb32 は、4・4・3・3行からなる14行の詩である。本作品においても、語末音であり、かつ脚韻の一部である子音 φt と φl がマークされており、 α 値が示す通り、この Topic 4 に含まれるキーワードが複数の作品に跨って現れていることが分かった。

4. 終わりに

本研究では、韻文作品の韻律を、本来、語の持つ潜在的意味を探索する目的で用いられる LDA トピックモデルによって分析可能であるかを検証するために、原文のままのテキスト、IPA 表記に変換したテキスト、そして IPA 表記を母音と子音に分割したテキストを対象コーパスとして各テキストに LDA を実行し、その結果を比較考察した。

分析及び考察の結果から、本研究で使用したデータに関しては、原文のままのテキストと IPA に変換したテキストでは、明確に英語詩の押韻を LDA が検知しているという確証は得られなかったが、IPA 表記を母音と子音に分割したテキストでは、他のテキストよりも押韻を検知していると見られる例が複数見つかった。一方で、IPA 表記のテキストを分割する単位を音節とした方がより良い結果を得られる可能性も示唆された。また、本稿では主に押韻の例として脚韻を挙げるに留めてしまったが、韻文作品で用いられる韻は脚韻のみではなく複数存在するため、より網羅的な韻の種類にも目を向けることが今後の課題である。

Tennyson の作品や、Tennyson の用いる音に関する研究は数多く存在しているが、多くが質的研究に依拠しているものである一方で、LDA あるいは他の計量的研究手法の利点の一つに、より多くのデータを客観的かつ網羅的に分析できるという点が挙げられる。Tennyson、ひいては他の韻文作品や文学作品の音に関するこれまでの先行研究の知見と融合させ、新たな知見を得るための計量的研究手法の可能性を本研究では提示した。

文献

- Blei, M. D., Ng, Y. A., and Jordan, I. M. (2003). “Latent Dirichlet Allocation.” *Journal of Machine Learning Research* 3, 2003, 993–1022.
- Fujita, I. (2021). “‘Were not Tennyson’s words delicious?’: An Analysis Using LDA Topic Modelling.” WEDHIA 2021, National Chengchi University.
- Henrichs, A. (2019). “Deforming Shakespeare’s Sonnets: Topic Models as Poems Author(s)” *Criticism*, 61 (3), 387–412.
- 黄 晨雯 (2020a). 「中国のミステリー小説におけるトピック解析の試み」, 『大阪大学言語文化学』, pp. 1–17.
- 黄 晨雯 (2020b). 「中国のミステリー小説をめぐる計量的分析—程小青と鬼馬星の作品を中心に—」, 『テキストマイニングとデジタルヒューマニティーズ 2019』 (大阪大学大学院言語文化共同研究プロジェクト2019) , pp. 31–45.

- Meeks, E. and Weingart, B. S. (2012). “The Digital Humanities Contribution to Topic Modeling.” *Journal of Digital Humanities*, 2 (1), 1–6.
- 黒田 絢香 (2017). 「小説テキストの計量的分析—アーサー・コナン・ドイルの作品から—」, 『テキストマイニングとデジタルヒューマニティーズ 2016』 (大阪大学大学院言語文化共同研究プロジェクト2016) , pp. 23–41.
- 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.
- 中尾 佳行・地村 彰之 (2016). 「『カンタベリー物語』の写本と刊本における言語と文体について」堀 正広 (編), 『コーパスと英語文体』 (pp. 21–52). 東京: ひつじ書房.
- Navarro-Colorado, B. (2018). “On Poetic Topic Modeling: Extracting Themes and Motifs From a Corpus of Spanish Poetry.” *frontiers in Digital Humanities*, (20), 5–15.
- Okabe, M. (2019). “Thou and You in Emily Dickinson's Poems Using Topic Modeling: Reconsideration of Interjections.” *Proceedings of Japanese Association for Digital Humanities Conference 2019*, 125–131.
- Rhody, M.L. (2012). “Topic Modeling and Figurative Language.” *CUNY Academic Works*, 2012, 19–35.
- 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.
- Tartakovsky, R. and Yehsyahu, S. (2019). “Meek as milk and large as logic: A corpus study of the non-standard poetic simile.” *Language and Literature*, 2019, 28 (3). 203–220.

MALLET 2.0.8: <https://mimno.github.io/Mallet/index> (最終閲覧日：2022年1月20日)

toPhonetics: <https://tophonetics.com/> (最終閲覧日：2022年1月20日)