Brian Powell

Dylan Goes Up On Trial:
Comparing Bob Dylan's Song Lyrics to Prize-Winning Poetry

Every October, the Swedish Academy announces the next recipient of its famed Nobel Prize in Literature. As nominations are only publicized 50 years later[1] and the selection process is nebulous at best, writers are resigned to the annual uncertainty surrounding the committee's decision. Who merits to join the ranks of literary giants such as W.B. Yeats, William Faulkner, Albert Camus, Gabriel García Márquez, and Toni Morrison?[2] The inherent subjectivity in determining the most worthy candidate inevitably leads to controversy among authors, poets, and critics alike. This discourse, usually confined to high-brow circles, garnered unprecedented mainstream attention in 2016 when the Swedish Academy named Bob Dylan its newest laureate, the first songwriter to receive the prestigious distinction. This ground-breaking selection prompted the question: should song lyrics be considered literature? And, if so, does Dylan's work rival the output of previous winners?

The Swedish Academy's Horace Engdahl presented the committee's perspective: "What brings about the great shifts in the world of literature? Often it is when someone seizes upon a simple, overlooked form, discounted as art in the high sense, and makes it mutate."[3] Thus Dylan was being recognized for his innovation, disseminating poetry to the masses through voice, not print, his words often vying for listeners' attention with an acoustic – and later electric – guitar. The Academy's Sara Danius added that the songwriter was by no means the first poet whose work was intended for performance, citing Homer and Sappho as predecessors of similar ilk.[4] Eminent novelist Salman Rushdie agreed, proclaiming Dylan as a "brilliant inheritor of the bardic tradition."[5]

Other writers were incensed, either disputing the literary merits of song lyrics or berating the selection committee for choosing a celebrity instead of a more deserving, lesser known candidate. *New York Times* editor Anna North argued that Dylan had already received countless awards in his own field, and that this crossover was one-way, since no novelist would ever win a Grammy.[6] Satirical writer Gary Shteyngart retorted: "I totally get the Nobel committee[,] [r]eading books is hard,"[7] insisting that the gulf in cultural capital between poetry and pop music is far from accidental. Author Jody Rosen was of the same mind, asserting that Dylan's lyrics are inseparable from his music as his "words don't sit inert on a page."[8]

While Rosen's statement may be true of concert-goers and other listeners, it does not apply to computational analyses of song lyrics, which isolate words from music. Using this approach, this report will compare Bob Dylan's songs to 16 poets' works in order to investigate whether his diction and themes distinguish him from other prize-winning bards. Moreover, if this difference exists, what lexical features make Dylan unique?

Over the course of his illustrious career, Dylan has played 716 songs live, for a total of 58,711 renditions.[9] However, his concerts are not representative of his output, because much of his work has never been performed and covers make up over half of those 716 pieces. Relying on data from Dylan's website,[10] I extracted lyrics for 395 songs – all of which were written by Dylan himself, some unreleased or only appearing on official bootlegs. I did not include *Tarantula*, his collection of experimental prose poetry written in 1966, since the Nobel Prize committee was recognizing his work "within the great American song tradition."[11]

While Dylan admits he has never paused to consider whether his songs should be considered literature,[12] his lyricism has inspired writers and critics to analyze his songs alongside celebrated poems. For example, the 2006 edition of *The Oxford Book of American Poetry*

featured *Highway 61 Revisited*'s "Desolation Row," and *The Cambridge Companion to Bob Dylan* was published in 2009.[13] Nevertheless, his connections to poets goes beyond these formal demarcations in part due to his close friendship with Allen Ginsberg – who notably appears in the music video to "Subterranean Homesick Blues" – and subsequent introduction to other Beat Generation writers. Dylan described his experience with these fellow icons of counter-culture, including Jack Kerouac and Lawrence Ferlinghetti, as "magic… it had just as big an impact on me as Elvis Presley."[14] The admiration was reciprocal: Ginsberg supposedly wept when he first heard "A Hard Rain's A-Gonna Fall" – although he maintained "Desolation Row" was his favourite Dylan song – believing Dylan to be a natural successor to his Beats.[15]

While Dylan and Ginsberg's social interactions are fascinating, did this influence impact the contents of their work? Furthermore, do any of these possible similarities extend beyond the Beat movement to other eminent poets? To investigate these questions, I analyzed 7,877 texts from sixteen 20th and 21st century poets,[16] including three Nobel laureates (Yeats, Eliot, Heaney), five Pulitzer Prize winners (Millay, Auden, Stevens, Williams, Plath), and Ginsberg himself. These documents were retrieved from .txtLAB's English poetry collection, and cover texts spanning the entirety of the authors' repertoires. I limited the study to 20th century figures since most were active during Dylan's career and would thus be exposed to similar cultural and literary shifts.

The breakdown by number of words and documents was the following:

	Bob Dylan	*Average Poet*	*Total (Combined)*
Documents	395	492	8,272
Total Words	113,796	101,282	1,734,305
Mean Length	288	206	210

Note that the total number of words in each corpus was similar, although Dylan's songs had 40% more terms per document than the poets' works. As an aside, while Dylan's work features choruses sparingly, these would engender more repetition and thus decrease the amount of unique words in a document.

In order to investigate how Dylan compares to renowned poets using text mining techniques, I generated similarity matrices using cosine similarity, either adjusting the sparsity or the presence of stopwords. I then explored the words that set Dylan apart from his poetic peers using Dunning's Log Likelihood and Fisher's Test.

Mean Similarity

I reasoned that creating a similarity matrix would be an effective first method of comparison. It would offer a quantitative measure of the contrast between the two collections, and, by varying the sparsity and presence of stopwords, indicate which types of words were responsible for the possible differences. For each trial, I used a vector space model to compute a cosine similarity score (as a percentage) between each of the 8,272 documents.

My first attempt was completed under the conditions of stemming, keeping stopwords, and requiring terms to appear in at least 5% of texts. These specifications were met by 440 words, and their raw counts were then scaled according to document length. The distribution of these frequencies within each work was used to generate similarity scores.

Subsetting the similarity matrix to examine Dylan songs' scores relative to the poems led to several noteworthy observations. The Dylan work with the lowest mean similarity (0.0825) was "Everything Is Broken" from his 1989 album *Oh Mercy*. The term "broken" accounts for almost one quarter of its 139 words, which would explain its lack of similarity to other texts.

On the other end of the spectrum, the song with the highest mean similarity score (0.6308) was somewhat unsurprisingly Ginsberg's favourite, "Desolation Row." This surreal account of an eclectic cast of characters – many of them related to literature, including Romeo, Cinderella, Ezra Pound, and T.S. Eliot – is often cited as Dylan's most poetic piece. While this score is indifferent to metaphor or imagery, clearly the song's diction contributes to its elevated literary recognition. Figures 1a and 1b show the striking contrast between the distribution of scores for the two songs.

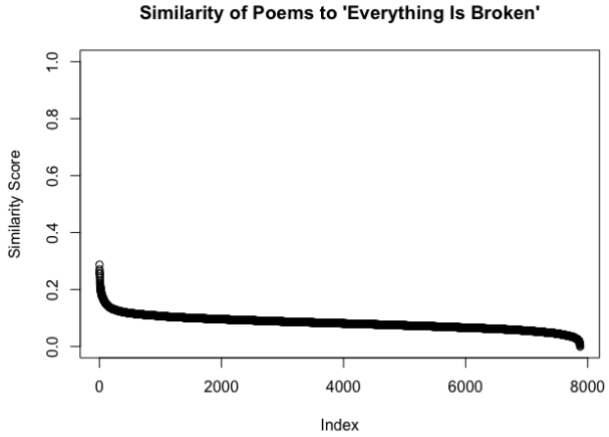

Fig. 1a: Decreasing similarity scores to Dylan's "Everything Is Broken"

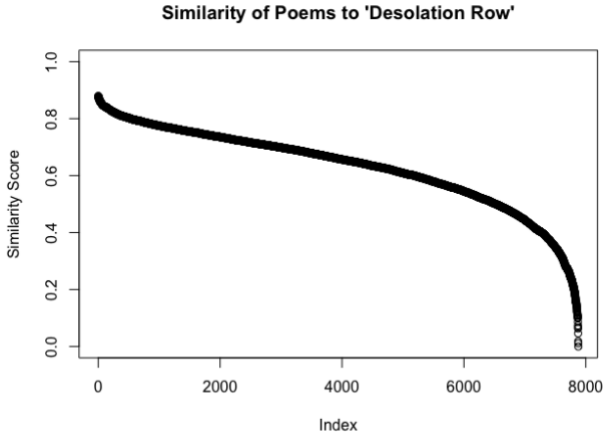

Fig. 1b: Decreasing similarity scores to Dylan's "Desolation Row"

Another Dylan song routinely quoted by his supporters in defense of his Nobel Prize is "Visions of Johanna," which received the second highest mean similarity with a score of 0.6306. Andrew Motion, former Poet Laureate of the United Kingdom, declared this *Blonde on Blonde* track to be the "best song lyric ever written,"[17] perhaps in part due to the correspondence between its word frequency distribution and that of other poems (although Motion may not describe it that way!).

While analyzing the similarity score extremes may add a quantitative argument to debates over Dylan's most poetic song, his entire repertoire must be considered when determining where

his songs fit within the scope of the other 7,877 poems. To generate this distribution, I compared each of the 16 poets pairwise by computing the mean score of 200 randomly selected entries from the subset of the similarity matrix with one poet's texts as rows and the other's as columns. This process resulted in 120 means, from which I re-sampled 16 values 1,000 times using bootstrapping and computed their average. A sample size of 16 was chosen to match the number of comparisons to Dylan's corpus. These 1000 means formed an approximately normal distribution (Shapiro-Wilk Normality test yielded $p = 0.1964 > 0.05$) with a mean of 0.492 and a standard deviation of 0.0162. To determine where Dylan's scores would fall on this distribution, he was compared to each of the other 16 poets using the same method as above, although without bootstrapping. Figure 2a shows the mean of these 16 average similarity scores superimposed on the poets' curve. As Dylan's score (0.434) is more than 3 standard deviations away from the mean, it can be considered an outlier relative to the poets' pairwise comparisons.

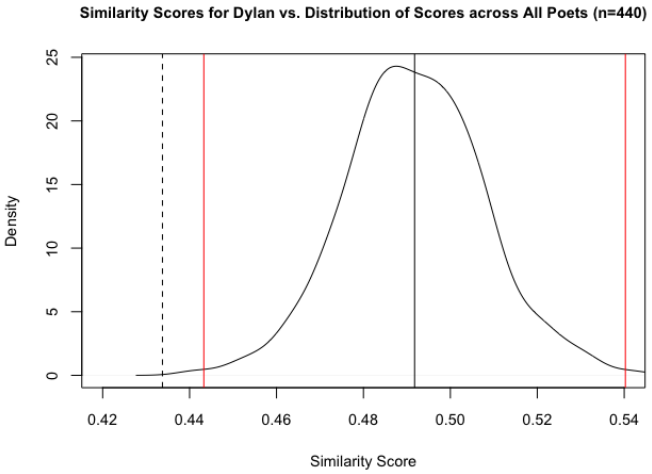

Fig. 2a: Distribution of average similarity scores across all poets for n=440 case; solid line is mean; red lines are delimiters for 3 standard deviations from the mean; dotted line is Dylan's average score

This observation suggests that Dylan's diction is distinctive from the poets' output, although I proceeded to conduct further trials to see if this result would be corroborated under other conditions. For the next trial, I adjusted the sparsity, instead requiring words to appear in 10% of texts. A new similarity matrix was created using the 202 terms that remained, and I carried out the same procedure as above to generate a distribution of average similarity scores with a mean of 0.523 and standard deviation of 0.0147. As shown in Figure 2b, Dylan's mean

score of 0.448 was even more standard deviations away from the mean than it was in the first trial. This suggests that the poets are more similar and Dylan more distinctive when considering the most commonly used words.

To further investigate this claim, I removed stopwords for my third trial and again required that terms appear in at least 5% of texts, such that 336 words remained. These conditions generated a distribution unlike the previous two (see Figure 2c), with a mean of 0.1030 and standard deviation of 0.00342. In this case, Dylan's average score (0.1031) was slightly above the mean, indicating that his texts were more similar to the group than the average poet's. Thus, when the most (and least) common terms are removed, Dylan fits in among this circle of poets. However, these scores are almost 5 times lower than in the previous two cases, which calls into question whether Dylan should be considered similar given that a 10% score, although on par with the other poets, does not provide strong evidence of similarity and could extend to a collection of songs by any artist.

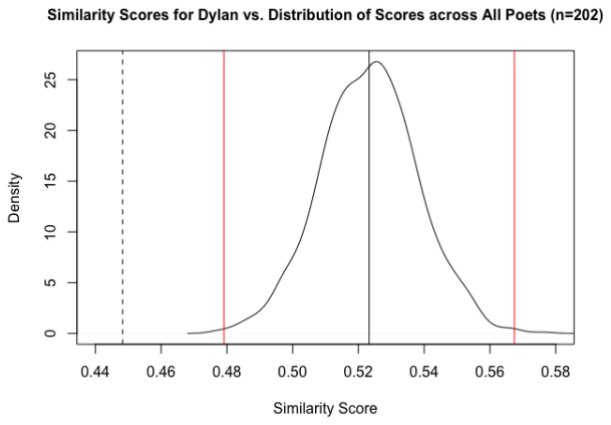

Fig. 2b: Distribution of average similarity scores across all poets for n=202 case

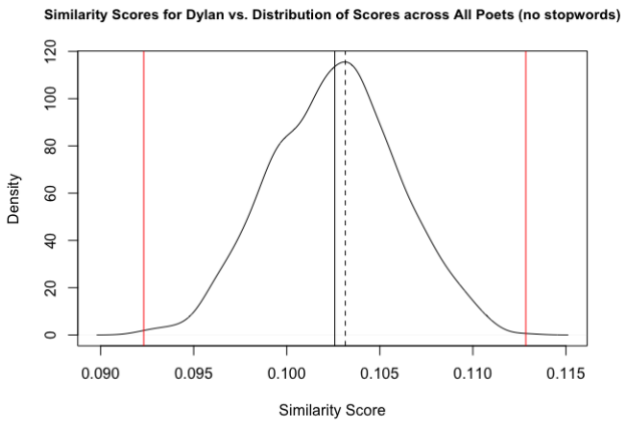

Fig. 2c: Distribution of average similarity scores across all poets for n=336 case (no stopwords)

These are numerous other reasons why these results should be interpreted with caution. Since I sampled from a set of means using bootstrapping, and then computed the mean of these 1,000 averages, the population distribution experienced significant flattening. As a consequence,

the standard deviation may have been reduced to the point where Dylan was unjustifiably considered an outlier. Moreover, Dylan's 16 comparisons were not enough to confidently establish the variance between his pairwise scores. Studying a greater number of poets would provide a distribution of Dylan mean similarity scores which could then be compared to the overall distribution. Lastly, one may question whether the top "n" words in a collection are what distinguish two poets, or whether their unique vocabulary is more representative of their work.

Distinctive Words

In the previous section, I remarked that the use of stopwords seemed to significantly influence two poets' similarity, and could be the feature that distinguishes Bob Dylan from the other writers. To test these hypotheses, I used Fisher's Exact Test and Dunning's Log-Likelihood Ratio on the top 3,000 terms to identify the distinctive words in Dylan's songs relative to the collection of poems. I did not opt for a Wilcoxon Rank Sum Test since the documents were short and should be considered as a group.

Since Dunning's test favours more frequent words, it would pinpoint the terms that I suspect engendered the differences in similarity observed in the first two trials. These G^2 values (see Figure 3) showed that Dylan was more likely to use singular personal pronouns (you, im, i, your, ill) than his poet counterparts, who preferred plural pronouns (our, we, us). This observation suggests that Dylan's songs have a different point of view than typical poems, usually juxtaposing the narrator (i, im, me, ill) – most often Dylan himself – with his intended audience (you, your,

word	G2_Sort.v
you	1134.33102
im	580.16062
gonna	485.02398
i	482.51405
well	473.08210
your	404.24815
got	394.13938
just	368.07304
dont	355.67654
aint	348.69528
me	327.28422
oh	243.97319
babi	213.44677
cant	195.94229
ill	151.05312

word	G2_Sort.v
of	-521.82539
the	-486.17516
our	-257.32269
as	-148.42943
we	-139.75693
their	-127.36114
green	-102.40269
at	-95.39570
us	-92.65210
white	-83.07085
and	-82.04841
which	-81.62150
is	-79.23188
in	-73.31828
thru	-66.72039

Fig. 3: G^2 values with significant words in Dylan songs (left) and poems (right)

occasionally babi) – namely a woman ("Just Like a Woman"), his critics ("Positively Fourth Street"), or society at large ("Gotta Serve Somebody"). He also uses contractions (im, dont, aint, cant, ill) at a significantly higher rate than the poets, who are more likely to employ definite articles, prepositions, and conjunctions (of, the, as, at, and, in). This contrast underscores the difference in formality between the two mediums. Interestingly, these observations match the conclusions from the comparison of pop songs to poems in Assignment 2, and further tests could examine if Dylan has more in common with other songwriters or these prize-winning poets.

Next, I used Fisher's Exact Test to identify unique words that appear significantly more often in one corpus than the other. As shown in Figure 4, the unique terms with the highest likelihood of appearing in Dylan's songs (and not the poems) were mostly slang words, often present participles with the 'g' removed (walkin, lookin, comin, mornin, nothin, talkin). This difference may stem from the simple fact that Dylan's writing is sung, and so his lyrics were transcribed based on his articulation. Since the poets' works are primarily consumed by readers, they are less likely to use this shortened spelling. Similarly, perhaps to reach a greater audience, Dylan avoids swearing or using racially-charged language, and so several of the distinctive words in the poetry collection were vulgarities or terms associated with race. It is important to note that these types of words are not necessarily employed by all poets, since a handful of writers may be responsible for all occurrences of this feature. Lastly, poets were more likely to refer to their own craft or the arts (poetri, imag, art, poet, poem [OR = 0.0930], music [OR = 0.1249]) than Dylan, which suggests that they are more introspective or self-referential than the songwriter.

word	fisher.OR
til	Inf
ya	99.0903988
walkin	70.0449441
lookin	69.0436612
wiggl	58.0308082
wanna	32.3623994
comin	29.5164561
mornin	29.0083126
gonna	27.2327875
babe	23.5210387
huh	19.3450411
goin	19.2649405
nothin	18.3441864
yeah	18.0105858
talkin	15.0066336
henri	14.6751100

word	fisher.OR
ass	0.00000000
fresh	0.00000000
fuck	0.00000000
stretch	0.00000000
mere	0.00000000
thigh	0.00000000
ignor	0.00000000
poetri	0.00000000
nigger	0.00000000
shit	0.04346700
imag	0.04442400
negro	0.04544285
green	0.06659330
enter	0.06663720
america	0.06663720
art	0.07403919

Fig. 4: Odds ratios with significant words in Dylan songs (left) and poems (right)

Conclusion and Further Testing

The Swedish Academy's controversial decision to award Bob Dylan the 2016 Nobel Prize in Literature prompted many critics to compare his work to other prize-winning poets. While most of these arguments were qualitative and focused on only a small portion of his output, this study used computational methods to analyze all of Dylan's lyrics to ascertain their similarity to the renowned poets' diction.

Generating similarity matrices under varying conditions showed that Dylan's work is indeed lexically distinct from the other poems, although some songs, particularly "Desolation Row" and "Visions of Johanna," had higher mean similarities than the average document. The statistical procedure used to generate the distribution of means could certainly be improved to reduce the flattening of values, which may have decreased the variance enough to make Dylan an outlier. Further tests could be conducted on a larger number of words or without stemming. Moreover, an ideal corpus would consist only of poems written by Nobel laureates, although this data may be difficult to access.

Testing for distinctive words then provided some insight into what words distinguish Dylan from the poets, and I claimed that formality, vulgarity, and intended audience were plausible bases for these differences. Again, a more diverse set of poets may result in an alternative group of distinctive terms. Then again, the most appropriate comparison may have been between Dylan and the Beats, since they shared many of the same experiences and milieus. The lasting question: did Ginsberg see some of himself in "A Hard Rain's A-Gonna Fall"? Enough to bring him to tears?

[1] Retrieved from the Nobel Prize website nobelprize.org.
[2] See note 1.
[3] Vanessa Thorpe and David Connett, “Bob Dylan Nobel prize speech: this is 'truly beyond words',” in *The Guardian* (11 Dec 2016).
[4] Krishnadev Calamur, “Bob Dylan has been awarded the Nobel Prize in Literature,” in *The Atlantic* (13 Oct 2016).
[5] Stephanie Merry, “Reactions to Bob Dylan’s Nobel Prize: Shock, elation and concern for Philip Roth,” in *The Washington Post* (13 Oct 2016).
[6] Anna North, “Why Bob Dylan Shouldn’t Have Gotten a Nobel,” in *The New York Times* (13 Oct 2016).
[7] Alex Beam, “Bob Dylan’s Nobel lecture: Short, sweet and on time (just),” in *The Boston Globe* (12 Jun 2017).
[8] See note 4.
[9] Retrieved from Dylan’s official website bobdylan.com.
[10] See note 9.
[11] See note 1.
[12] See note 3.
[13] Ben Sisario, Alexandra Alter, and Sewell Chan, “Bob Dylan Wins Nobel Prize, Redefining Boundaries of Literature,” in *The New York Times* (13 Oct 2016).
[14] Sean Wilentz, “Bob Dylan, the Beat Generation, and Allen Ginsberg’s America,” in *The New Yorker* (13 Aug 2010).
[15] See note 14.
[16] The full list consists of: William Butler Yeats, T.S. Eliot, Edna St. Vincent Millay, W.H. Auden, Sylvia Plath, Wallace Stevens, Langston Hughes, Muriel Rukeyser, William Carlos Williams, Philip Larkin, Allen Ginsberg, Adrienne Rich, Seamus Heaney, Amiri Baraka, Al Young, and Wanda Coleman.
[17] Vanessa Thorpe, “Laureate gives laurels to Dylan,” in *The Guardian* (Oct 3 1999).