BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

Show Up and *Turn Up* in American English and British English: A Multi-Method Analysis Using COCA and BNC

Namkil Kang [1], ***Eun Hee Kim*** [2*]

[1]*College of Hotel Management, Far East University, Gamgok-myeon, Eumseong-gun 27621, Republic of Korea*

[2]*College of Liberal Arts, Far East University, Gamgok-myeon, Eumseong-gun 27621, Republic of Korea*

ABSTRACT

This article investigates the relationship between the terms *show up* and *turn up* in the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). In the COCA, the terms exhibit a 25% similarity in ranking, while in the BNC, this similarity is 14.28%. In terms of genre-specific usage, *show up* and *turn up* are most distant in the fiction genre in the COCA, while in the BNC, they are furthest apart in the spoken genre. The closest similarity occurs in the TV/movie genre for the COCA and in the magazine genre for the BNC. Frequency analysis reveals significant national variation. In American English, *show up* shows greater fluctuation from the mean, while *turn up* displays more consistent frequency. In contrast, British English usage demonstrates a more stable frequency for *show up*, while *turn up* exhibits more erratic variation. The standard deviations for *show up* in the COCA (1,134) and the BNC (18) further highlight this disparity, as *turn up* frequencies in both corpora show opposite trends. Statistically, the COCA reveals a strong positive correlation (r = 0.7375) between the two terms, suggesting a significant relationship in American English. However, the BNC's correlation coefficient (r = 0.0669) indicates no meaningful connection between the terms in British English. This comparison underscores notable national variations in the usage and relationship between *show up* and *turn up* across the two varieties of English.

Keywords: Ranking; Euclidean Distance; Correlation; Standard Deviation; Linear Regression

*CORRESPONDING AUTHOR:

Eun Hee Kim, College of Liberal Arts, Far East University, Gamgok-myeon, Eumseong-gun 27621, Republic of Korea; Email: ekim@kdu.ac.kr

ARTICLE INFO

Received: 11 April 2025 | Revised: 24 April 2025 | Accepted: 13 May 2025 | Published Online: 16 May 2025

DOI: https://doi.org/10.30564/fls.v7i5.9432

CITATION

Kang, N., Kim, E.H., 2025. Show up and Turn up in American English and British English: A Multi-Method Analysis Using COCA and BNC. Forum for Linguistic Studies. 7(5): 1002–1013. DOI: https://doi.org/10.30564/fls.v7i5.9432

1. Introduction

This article aims to evaluate whether there is a significant relationship between the terms *show up* and *turn up* in the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). A specific goal of this research is to examine whether national variation exists in the usage of these terms across American and British English. The research was conducted using Python, a powerful tool for data analysis. In contemporary English, phrasal verbs such as *show up* and *turn up* are frequently used across both spoken and written discourse, playing an essential role in everyday communication. These two expressions are often regarded as near-synonyms and are commonly used interchangeably in a wide range of contexts to convey the idea of arrival or appearance. Despite their frequent usage and apparent semantic overlap, there has been surprisingly little systematic linguistic inquiry into their nuanced distinctions, contextual preferences, or regional variations. This gap in the literature highlights the importance and relevance of the present study, which seeks to explore the subtle differences between *show up* and *turn up*, as well as their distributional patterns in American and British English. The primary corpora used for this investigation are the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC), both of which are well-established resources in the field of corpus linguistics. COCA is widely recognized as one of the most comprehensive corpora available for American English [1]. It contains over one billion words drawn from eight major genres, including spoken, fiction, magazines, newspapers, academic texts, blogs, web content, and television/movies. This corpus offers an extensive and balanced dataset that supports robust linguistic analysis. The frequency data for *show up* and *turn up* in COCA are publicly accessible and were retrieved via Google in March 2025. In contrast, the British National Corpus (BNC) [2], originally compiled by Oxford University Press in the late 1980s and early 1990s, contains approximately 100 million words of text from a wide array of genres, including spoken language, fiction, newspapers, academic writing, and other written forms. It serves as a foundational resource for studies of British English. Like COCA, the BNC also provides frequency data for *show up* and *turn up*, which remains publicly available online. By analyzing data from both corpora, this study offers a cross-regional perspective on how these expressions are used and understood in different varieties of English, contributing to a more refined understanding of phrasal verb usage in contemporary language. To begin with, we will conduct a ranking analysis (as defined in our study), which involves assessing the overall frequency of *show up* and *turn up* across eight genres in the COCA and seven genres in the BNC. This analysis will present the rankings of each genre in descending order, allowing us to assess the similarity between *show up* and *turn up* in both corpora. Next, we will calculate the Euclidean distance between *show up* and *turn up* in each genre. The Euclidean distance serves as a measure of similarity—the smaller the distance, the greater the similarity between the two terms. Additionally, we will perform variance and standard deviation analyses to better understand the distribution of *show up* and *turn up* within the COCA and BNC. These analyses will help us explore how frequently each term appears and whether their distribution is consistent across genres. Gries (2006) discusses methodological considerations for identifying and quantifying variation in corpus-linguistic studies [3]. While not centered on specific phrasal verbs, the study emphasizes the importance of statistical measures like variance and standard deviation in understanding linguistic variability across different corpora. Following this, we will conduct a correlation analysis to assess the degree of correlation between *show up* and *turn up* in both corpora. This will provide insights into how closely these two terms are related in terms of their usage patterns. Further, a linear regression analysis will be carried out to explore how much the independent variable *show up* influences the dependent variable *turn up*. By examining the results of the regression analysis, we will gain a better understanding of how the frequency of *show up* might impact the frequency of *turn up*. Finally, we will perform a Chi-Squared test analysis, a standard statistical method used to determine whether there is a significant association between two categorical variables—specifically, between *show up* and *turn up*. Through these various analyses, this article aims to provide a comprehensive examination of the usage patterns of *show up* and *turn up* in both American and British English, shedding light on any potential linguistic variations and associations between the two terms. For further study of corpus

linguistics, refer to the following references and corpora: Aarts and Granger (1993), Baker, Hardie, and McEnery (2006), Barlow and Kuperman (2008), Bier (1993), Biber, Conrad, and Reppen (1998), Channel (2000), Firth (1957), Gries (2006, 2013), Hunston and Francis (2000), McEnery and Wilson (2001), Meyer (2002), O'Keeffe, McCarthy, and Carter (2007), Sinclair (1991), and Tognini-Bonelli (2001) and the Corpus of Historical American English, and the Hansard Corpus [4-19].

2. Materials and Methods

This article presents a detailed and methodologically rigorous analysis of the usage patterns of the expressions *show up* and *turn up* within two major English corpora: the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). The COCA comprises eight distinct genres, while the BNC encompasses seven, offering a broad representation of linguistic variation across both American and British English. Using Python, we processed frequency data for *show up* and *turn up* across all available genres in each corpus. These datasets, which remain publicly accessible via Google, form the foundation of our comparative study. Our analysis employed a range of quantitative techniques to capture both frequency trends and deeper distributional patterns. Specifically, we conducted a ranking analysis (as defined in section 4.1.), Euclidean distance analysis to assess lexical proximity, variance and standard deviation analyses to measure dispersion, correlation and linear regression analyses to identify potential predictive relationships, and Chi-Squared tests to evaluate the statistical significance of observed differences. Taken together, these analytical approaches enable us to investigate the nuanced relationship between *show up* and *turn up*, while also shedding light on regional linguistic preferences and variation between American and British English usage.

3. Data Collection

In March 2025, the frequencies of the verb phrases *show up* and *turn up* were systematically extracted from two major English language corpora: the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). The COCA comprises a total of eight distinct genres, while the BNC includes seven, each representing a broad spectrum of written and spoken English usage. For each expression, frequency data were collected across all genres within both corpora to ensure a balanced and comprehensive comparative analysis. These frequency distributions served as the foundation for seven distinct types of analysis, each designed to investigate different dimensions of usage patterns and linguistic variability. The analyses included ranking analysis, Euclidean distance calculation, variance and standard deviation measures, correlation and linear regression modeling, and Chi-Squared testing, all of which were conducted using Python-based statistical tools. To provide a clear overview of the raw data, **Tables 1** and **2** present the frequencies of *show up* and *turn up* across the eight genres of the COCA and the seven genres of the BNC, respectively. These tables serve as essential reference points for understanding the patterns

Table 1. Ranking Analysis: COCA.

Ranking	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8
Show up (Frequency)	TV/Movie (4,695)	Blog (3,470)	Spoken (3,213)	Web (2,519)	Mag (2,479)	News (2,402)	Fic (2,134)	Acad (472)
Turn up (Frequency)	TV/Movie (928)	Fic (676)	Mag (650)	Blog (455)	News (441)	Spoken (431)	Web (394)	Acad (169)

Table 2. Ranking Analysis: BNC.

Ranking	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7
Show up (Frequency)	Fic, Misc (68) (68)		Non-acad (59)	Mag (46)	Acad (45)	Spoken (31)	News (21)
Turn up (Frequency)	Fic (267)	Spoken (223)	Misc (175)	News (174)	Mag (110)	Non-acad (80)	Acad (42)

identified in our subsequent quantitative analyses.

4. Results

4.1. Ranking Analysis

In the following section, we seek to compare the rankings of the terms *show up* and *turn up* across the eight genres of the COCA and the seven genres of the BNC. The term ranking analysis refers to the process of evaluating and comparing the overall frequency order of these terms across each genre within both corpora. This analysis allows us to assess how the terms are distributed and used in different contexts. To conduct this analysis, we gathered the frequency data for *show up* and *turn up* from each genre in both the COCA and the BNC in March 2025. By examining these rankings, we can gain insights into the relative prominence of each term within specific genres and make comparisons between the two corpora.

It is important to note that the overall frequency of *show up* is 21,384, whereas *turn up* occurs 2,134 times in the COCA. This significant difference indicates a clear preference for *show up* over *turn up* in American English. However, it is equally crucial to highlight that in the TV/movie genre, both *show up* and *turn up* exhibit the highest frequency and proportion. In contrast, in the academic genre, these terms have the lowest frequency and proportion. From these observations, it is apparent that *show up* and *turn up* are most commonly used in the TV/movie genre, while they are least utilized in the academic genre. It is also worth noting that the TV/movie genre is the most influenced by *show up*, followed closely by the blog genre, spoken genre, web genre, magazine genre, newspaper genre, fiction genre, and finally, the academic genre, in that order. On the other hand, the TV/movie genre is also the most influenced by *turn up*, followed in descending order by the fiction genre, magazine genre, blog genre, newspaper genre, spoken genre, web genre, and academic genre. Importantly, *show up* and *turn up* share identical rankings in two genres—the TV/movie and academic genres—while they display different rankings in the other six genres. This suggests that the two terms are approximately 25% similar, indicating a relatively low degree of similarity. Therefore, while *show up* and *turn up* can be considered synonyms, their low degree of similarity across genres implies nuanced differences in usage. Now, turning our attention to the BNC, let us examine **Table 2**, which illustrates the rankings of *show up* and *turn up* across the seven genres.

It is worth noting that, as illustrated in **Table 2**, *show up* and *turn up* exhibit the highest frequency and proportion in the fiction genre in the BNC. This suggests that, within the UK, *show up* and *turn up* are most frequently used in the fiction genre. Additionally, it is important to highlight that in the miscellaneous genre, *show up* also has the highest frequency, similar to its prominence in the fiction genre. This indicates that, among the seven genres, *show up* occurs most frequently in both the fiction and miscellaneous genres. Conversely, in the newspaper and academic genres, *show up* and *turn up* demonstrate the lowest frequency and proportion, respectively. This suggests that these two expressions are less commonly used in these particular genres. As shown in **Table 2**, the fiction and miscellaneous genres are most strongly influenced by *show up*, followed by the non-academic genre, the magazine genre, the academic genre, the spoken genre, and, finally, the newspaper genre, in descending order. On the other hand, the fiction genre is the most influenced by *turn up*, followed by the spoken genre, the miscellaneous genre, the newspaper genre, the magazine genre, the non-academic genre, and the academic genre, in that order. Most importantly, *show up* and *turn up* share the same ranking in the fiction genre, yet they display different rankings in the remaining six genres. This suggests that the two terms are 14.28% similar in their ranking analysis across the British corpus. It is also noteworthy that in American English, *show up* and *turn up* are 25% similar, while in British English, this similarity drops to 14.28%. This disparity indicates a small but notable national variation between American and British English in the usage of these terms.

4.2. Euclidean Distance Analysis

This section is dedicated to exploring the similarity between the terms *show up* and *turn up* across each genre, using Euclidean distance as a measure of similarity. Euclidean distance serves as a quantitative indicator of how closely the two terms are related in each genre. The closer the distance between them, the higher the degree of similarity between *show up* and *turn up*. In other words, a smaller Euclidean distance suggests that the terms are

more alike in their usage within a particular genre. To quantify this relationship, we define the Euclidean distance in the following manner:

$$\sqrt{(p1-q1)2+(p2-q2)2+(pn-qn)2} \quad (1)$$

Now, let us examine **Table 3**, which displays the Euclidean distance between *show up* and *turn up* across each genre. To calculate this, we converted the frequency of *show up* and *turn up* in each genre into their respective proportions within each genre.

It is important to highlight that, as shown in **Table 3**, *show up* is the furthest from *turn up* in the fiction genre. Specifically, the Euclidean distance between the two terms in this genre is 6.34, the highest recorded distance across all genres. This suggests that *show up* and *turn up* exhibit the lowest degree of similarity in the fiction genre. On the other hand, the fiction genre is followed closely by the blog genre. As indicated in **Table 3**, *show up* is the second furthest from *turn up* in the blog genre, with a distance that implies the second lowest similarity between the two terms in this genre. In contrast, it is worth noting that *show up* is the closest to *turn up* in the TV/movie genre. In this genre, the Euclidean distance between the two terms is 0.44, the lowest recorded in the analysis. This strongly suggests that *show up* and *turn up* share the highest degree of similarity within the TV/movie genre. Interestingly, *show up* is the second closest to *turn up* in the newspaper genre, where the Euclidean distance is 0.59, making it the second smallest. This indicates that *show up* and *turn up* exhibit the second highest similarity in the newspaper genre. From these observations, it is reasonable to conclude that *show up* is the furthest from *turn up* in the fiction genre, while it is closest to *turn up* in the TV/movie genre, highlighting the varying degrees of similarity between the two terms across different genres. Now, let us shift our focus to the BNC. Refer to **Table 4** for the relevant data:

It is intriguing to observe, as illustrated in **Table 4**, that *show up* is the furthest from *turn up* in the spoken genre. In this genre, the Euclidean distance between the two terms is 11.65, which represents the highest distance observed across all genres. This suggests that *show up* and *turn up* exhibit the lowest degree of similarity in the spoken genre. Additionally, it is noteworthy that *show up* is the second furthest from *turn up* in the newspaper genre, which implies that these terms show the second lowest level of similarity within this genre. Conversely, *show up* is closest to *turn up* in the magazine genre, with an actual distance of 3.33, the lowest distance recorded. This suggests that the two terms share the highest degree of similarity in the magazine genre. Following closely behind, the miscellaneous genre exhibits the second smallest Euclidean distance, with a value of 3.78. This indicates that *show up* and *turn up* display the second highest similarity in the miscellaneous genre. To summarize, *show up* is the furthest from *turn up* in the spoken genre, while it is closest to *turn up* in the magazine genre. For the COCA, *show up* is the furthest from *turn up* in the fiction genre, whereas the two terms are closest in the TV/movie genre. On the other hand, in the BNC, *show up* and *turn up* are most distant in the spoken genre, while they are closest in the magazine genre. This comparison highlights a notable national variation between American English and British English, indicating that the two corpora exhibit different patterns of similarity between *show up* and *turn up* across various genres.

Table 3. Euclidean Distance: COCA.

Genre	Blog	Web	TV/M	Spoken	Fic	Mag	News	Acad
Show up	16.22	11.77	21.95	15.02	9.97	11.59	11.23	2.20
Turn up	10.97	9.50	22.39	10.47	16.31	15.68	10.64	4.07
Euclidean Distance	5.25	2.27	0.44	4.55	6.34	4.09	0.59	1.87

Table 4. Euclidean Distance: BNC.

Genre	Spoken	Fic	Mag	News	Non-acad	Acad	Misc
Show up	9.17	20.11	13.60	6.21	17.45	13.31	20.11
Turn up	20.82	24.92	10.27	16.24	7.46	3.92	16.33
Euclidean Distance	11.65	4.81	3.33	10.03	9.99	9.39	3.78

4.3. Variance Analysis

The goal of this section is to examine the distribution of *show up* and *turn up* in both the COCA and the BNC. We aim to compare the distribution of *show up* with that of *turn up* and evaluate their respective patterns. Additionally, we seek to assess whether there is any national variation between American English and British English in their usage. To begin, let us examine the mean and variance of *show up* and *turn up* in the COCA (**Table 5**).

Table 5. The Variance of *Show Up* and *Turn Up*: COCA.

	Mean	Variance
Show up	2,673.0	1,285,626.0
Turn up	518.0	45,641.5

As shown in **Table 5**, the variance for *show up* is 1,285,626.0. This figure represents a measure of dispersion, indicating how spread out the data are around the mean. A higher variance suggests that the frequencies of *show up* are more widely scattered from their average value, meaning that the individual occurrences of *show up* tend to vary significantly. A variance of 1,285,626 is notably large, which indicates considerable fluctuations in the frequency of *show up* across the different genres. In contrast, the variance for *turn up* is 45,641.5. A lower variance indicates that the frequencies of *turn up* are more tightly clustered around the mean, showing less variation. With a variance of 45,641.5, the usage of *turn up* remains relatively stable, implying that its frequency fluctuates less compared to *show up*. Thus, the frequencies of *show up* display much greater variability than those of *turn up*, suggesting that *show up* is used in a more inconsistent or varied manner. This greater fluctuation in the data for *show up* points to a less predictable usage pattern compared to *turn up*. Now, let us turn our attention to the BNC. Refer to **Table 6** for the relevant data:

Table 6. The Variance of *Show Up* and *Turn Up*: BNC.

	Mean	Variance
Show up	48.285714285714285	325.23809523809524
Turn up	153.0	6386.666666666667

It is important to note that *show up* has a variance of 325.24, which indicates that its frequencies are relatively consistent around the mean of 48.29. This lower variance suggests a more stable usage pattern for *show up*, implying that its occurrences tend to stay within a predictable range. In contrast, *turn up* exhibits a much higher variance of 6,386.67, which reflects that its frequencies are more widely dispersed around its mean of 153.0. The higher variance of *turn up* indicates a more erratic usage pattern, with some instances of the term showing very high frequencies, while others occur much less frequently. To summarize, in American English, the frequencies of *show up* tend to vary greatly from the mean, while *turn up* frequencies are more tightly clustered around their average value. Conversely, in British English, *show up* frequencies exhibit a more consistent and stable usage pattern, whereas *turn up* frequencies are more erratic, with greater fluctuation from the mean. From these observations, it is clear that there is a noticeable national variation between American English and British English in terms of the variance of *show up* and *turn up*. This suggests that the usage patterns of these two expressions differ significantly between the two varieties of English, further highlighting the distinct linguistic characteristics found in each.

4.4. Standard Deviation Analysis

In the following analysis, we examine the standard deviation of the frequencies of *show up* and *turn up* in both the COCA and the BNC. Additionally, we compare the distribution of these phrases across the two corpora. Based on this comparison, we assess whether there is national variation between American English and British English. Let us now turn to **Table 7**:

Table 7. Standard Deviation: COCA.

	Mean	Variance	Standard Deviation
Show up	2,673.0	1,285,626.0	1,134
Turn up	518.0	45,641.5	214

It is worthwhile noting that the standard deviation of *show up* frequencies is 1,134, which signifies a spread of 1,134 around the mean value of 2,673. This means that the frequencies of *show up* generally fall within the range of approximately 1,539 to 3,807, providing insight into the variability of its usage within the dataset. A higher standard deviation suggests greater fluctuation in occurrence across different contexts. Similarly, the standard deviation of *turn up* frequencies is 214, indicating a spread of 214 around the mean value of 518. This implies that the frequencies of

turn up typically range between 304 and 732. Compared to *show up*, the smaller standard deviation suggests that *turn up* exhibits more consistent usage within the corpus. With this in mind, let us now shift our focus to the BNC (**Table 8**):

Table 8. Standard Deviation: BNC.

	Mean	Variance	Standard Deviation
Show up	48.285714285714285	325.23809523809524	18
Turn up	153.0	6386.666666666667	80

It is important to highlight that the standard deviation of *show up* in the BNC is 18, indicating a spread of 18 around the mean value of 48.286. In other words, the frequencies of *show up* generally fall within the range of approximately 30.286 to 66.286. This relatively small standard deviation suggests a more consistent usage of the phrase within the corpus. Conversely, the standard deviation of *turn up* in the BNC is significantly higher at 80, reflecting a spread of 80 around the mean value of 153. This means that the frequencies of *turn up* typically range between 73 and 233, indicating greater variation in its usage across different contexts. To summarize, in the COCA, *show up* has a standard deviation of 1,134 around a mean of 2,673, while *turn up* has a standard deviation of 214 around a mean of 518. In contrast, in the BNC, *show up* exhibits a much smaller spread of 18 around a mean of 48.286, whereas *turn up* has a more substantial spread of 80 around a mean of 153. This contrast reinforces the idea that *show up* and *turn up* demonstrate opposite trends in terms of variability in American and British English. Once again, the data highlights a clear national variation between the two varieties of English, particularly in the way these expressions fluctuate in frequency within their respective corpora.

4.5. Correlation Analysis

The objective of this section is to assess whether there is a correlation between the frequencies of *show up* and *turn up* in both the COCA and the BNC. By analyzing their relationship, we aim to determine whether the usage patterns of these expressions exhibit any statistical association within each corpus. Additionally, we seek to evaluate whether our findings align more closely with the null hypothesis or the alternative hypothesis. This analysis will help us determine whether any observed variations in frequency occur by chance or if there is a meaningful underlying pattern distinguishing the two phrases. With this in mind, let us now turn our attention to **Table 9**:

Table 9. Pearson Correlation Coefficient: COCA.

		Turn Up
Show up	*r*	0.7374913489249744
	p	0.03678785552557782
	N	8

The correlation coefficient, denoted as *r*, measures both the strength and direction of the linear relationship between the frequencies of *show up* and *turn up*. As presented in **Table 9**, the calculated *r* value is 0.7375, indicating a strong positive correlation between the occurrences of these two phrases. In other words, as the frequency of *show up* increases, the frequency of *turn up* tends to increase as well, suggesting a notable degree of association between their usage patterns. It is important to recall that the *r* value ranges from −1 to +1. A value close to +1 signifies a strong positive correlation, meaning that the two variables tend to increase together. Conversely, a value near −1 would indicate a strong negative correlation. Simply put, as one variable increases, the other decreases. A value around 0, on the other hand, would suggest little to no linear relationship between the variables. To better understand the strength of this correlation, let us now examine **Table 10**, which provides a detailed interpretation of different *r* values and their corresponding levels of correlation strength:

Table 10. Criterion for Correlation.

0.00 to 0.19	Very weak or no correlation
0.20 to 0.39	Weak correlation
0.40 to 0.59	Moderate correlation
0.60 to 0.79	Strong correlation
0.80 to 1.00	Very strong or perfect correlation

The strong positive correlation between the frequencies of *show up* and *turn up* suggests that as the occurrence of *show up* increases, the occurrence of *turn up* tends to increase as well. This indicates that the two expressions exhibit similar usage patterns within the corpus, reinforcing the idea that their distributions are not independent of each other. To further evaluate the significance of this correlation, we rely on the *p*-value, which is used to test the

null hypothesis. The *p*-value helps determine whether the observed relationship between the two variables is due to random chance or represents a meaningful statistical pattern. If the *p*-value is less than 0.05, we reject the null hypothesis, indicating that the correlation is statistically significant. As shown in **Table 9**, the *p*-value for this analysis is 0.037, which falls below the 0.05 threshold. This result confirms that the correlation between *show up* and *turn up* is statistically significant, providing strong evidence against the null hypothesis. In other words, the likelihood that this correlation has occurred by chance is minimal, reinforcing the idea that the two phrases are meaningfully associated in their usage patterns. With this in mind, let us now shift our focus to the BNC (**Table 11**):

Table 11. Pearson Correlation Coefficient: BNC.

		Turn Up
Show up	*r*	0.06695599740462654
	p	0.8865865899256176
	N	7

As shown in **Table 11**, the correlation coefficient (*r*) in the BNC is 0.0669. This low *r* value indicates that there is no meaningful correlation between the frequencies of *show up* and *turn up* in the BNC. Unlike in the COCA, where the two expressions exhibit a strong positive correlation, their usage in the BNC appears to be largely independent of one another. Interestingly, the slight positive *r* value suggests that as the frequency of *show up* increases, the frequency of *turn up* does not follow a consistent trend and may even decrease in certain instances. This lack of association implies that the relationship observed in American English does not hold in British English. Furthermore, the *p*-value for this analysis is 0.8866, which is much higher than the 0.05 cutoff for statistical significance. This means that we fail to reject the null hypothesis, confirming that the observed correlation in the BNC is not statistically significant and likely occurs due to random chance rather than an underlying linguistic pattern. In summary, while *show up* and *turn up* are strongly correlated in the COCA, no such relationship exists in the BNC. This stark contrast further highlights the national variation between American and British English, reinforcing the idea that these expressions behave differently in each variety of English.

4.6. Linear Regression Analysis

The aim of this section is to evaluate the extent to which the frequencies of *show up* influence the frequencies of *turn up*. To explore this relationship further, let us now examine **Table 12**:

The R^2 value indicates the proportion of variation in the frequency of *turn up* that can be explained by the frequency of *show up*. A high R^2 value, closer to 1, suggests that *show up* is a strong predictor of *turn up*, accounting for most of the variation in its usage. Conversely, a low R^2 value implies that *show up* does not significantly predict *turn up* and that other factors may be influencing the latter's frequency. In this case, the R^2 value of 0.5439 indicates that *show up* explains only about 54.39% of the variation in *turn up*. Since this value is not particularly close to 1, it suggests that while there is some level of association between the two expressions, *show up* does not strongly predict the frequency of *turn up*. With this in mind, let us now turn our attention to **Table 13**:

Table 12. Model Summary: COCA.

Model	**R**	**R^2**	**Adjusted R^2**	**Std. Error**
1	0.7374913489249744	0.5438934897391783	0.46787573802904137	166.6028622726084

Table 13. Coefficients.

Model	**B**	**Std. Error**	**Beta**	***t***	***p***
1 (Constant)	146.568681	150.837118	NaN	0.971702	0.368726
Show up	0.138957	0.051949	0.737491	2.674850	0.036788

If the *p*-value is less than 0.05, it indicates that the effect is statistically significant, meaning that *show up* has a meaningful influence on *turn up*. On the other hand, if the *p*-value is greater than 0.05, the effect is not considered statistically significant, suggesting that *show up* may not have a strong influence on *turn up*. As shown in **Table 13**, the *p*-value for this analysis is 0.0368, which is below the 0.05 threshold. This means that the effect is statistically significant, and we can conclude that *show up* does, in fact, have an impact on the frequency of *turn up*. The ANOVA results support this finding, confirming that the relationship between the two expressions is not due to random chance. With this in mind, let us now shift our focus to the BNC. Refer to **Tables 14** and **15** for further details:

As demonstrated in **Table 13**, the R^2 value, which is not close to 1, indicates that the frequency of *turn up* in the BNC is not significantly explained by the frequency of *show up*. This suggests that *show up* does not serve as a strong predictor for the variations in *turn up* within the BNC. Furthermore, the ANOVA results show that the *p*-value is greater than 0.05, which implies that the frequency of *show up* does not significantly affect the frequency of *turn up* in the BNC. Additionally, the standard error values presented in **Tables 14** and **15** provide an estimate of the accuracy of predictions made within the model. A higher standard error would indicate less accurate predictions, while a lower standard error would suggest more precise predictions. To summarize, in the COCA, the frequency of *show up* significantly influences the frequency of *turn up*, reflecting a clear relationship between the two expressions in American English. In contrast, the frequencies of *turn up* in the BNC are not explained by *show up* frequencies, highlighting a distinct difference between American and British English. This further underscores the notable national variation between the two varieties of English in terms of how these expressions are used and related.

4.7. Chi-Squared Test Analysis

The aim of this section is to evaluate whether there is a significant association between the *show up* frequencies and the *turn up* frequencies.

The Chi-Squared statistic (χ^2) measures the difference between observed and expected frequencies, allowing us to assess whether there is a significant relationship between two categorical variables—in this case, *show up* and *turn up*. A higher χ^2 value indicates a larger discrepancy between the observed and expected values, suggesting a potential relationship between the two variables. The expected values represent what we would anticipate if there were no association between *show up* and *turn up*; these are calculated under the assumption that the two variables are independent. If the observed frequencies differ substantially from the expected frequencies, this signals a meaningful relationship between the variables. The χ^2 statistic quantifies the magnitude of this difference. As illustrated in **Table 16**, the high χ^2 value of 353.59 indicates a considerable difference between the observed and expected frequencies, which strongly suggests that *show up* and *turn up* are related in some way. Additionally, the *p*-value of 2.10×10^{-72}, which is effectively 0, indicates that the probability of this result occurring by random chance is virtually nonexistent. Given that the *p*-value is much smaller than the 0.05 threshold, we can confidently conclude that there is a statistically significant relationship between *show up* and *turn up*. With this in mind, let us now turn our attention to the BNC (**Table 17**).

Table 14. Model Summary: BNC.

Model	R	R^2	Adjusted R^2	Std. Error
1	0.06695599740462803	0.004483105588448555	−0.19462027329386178	87.34781896973806

Table 15. Coefficients.

Model	B	Std. Error	Beta	*t*	*p*
1 (Constant)	138.673353	101.022883	NaN	1.372692	0.228218
Show up	0.296706	1.977314	0.066956	0.150055	0.886587

Table 16. Chi-Squared Test: COCA.

Chi-Squared Statistic	353.58908932717065
p-value	2.100634525637004e-72

Table 17. Chi-Squared Test: BNC.

Chi-Squared Statistic	107.92039802248735
p-value	5.5538440055220374e-21

The high Chi-Squared value of 107.92 indicates that the observed frequencies of *show up* and *turn up* differ significantly from the expected frequencies. This substantial difference suggests that the two expressions are not independent and likely exhibit a meaningful relationship. In other words, the frequencies of *show up* and *turn up* appear to be linked in some way. Furthermore, the *p*-value associated with this result is extremely small—much smaller than the 0.05 threshold for statistical significance. This indicates that the probability of this relationship occurring by chance is virtually zero, providing strong evidence against the null hypothesis, which assumes no relationship between *show up* and *turn up*. Given this, we confidently reject the null hypothesis in favor of the alternative hypothesis, which posits a significant relationship between the two expressions. To summarize, in the COCA, we observe a clear and statistically significant relationship between *show up* and *turn up*. Similarly, in the BNC, we also find evidence of a relationship between these two expressions, though this leads to a smaller degree of national variation between American English and British English. While both corpora suggest a connection, the nature of the relationship appears to differ slightly across the two varieties of English, pointing to some degree of variation between them.

5. Discussion

In the following analysis, we examine the relationship between the terms *show up* and *turn up* within two prominent corpora: the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). Notably, in the COCA, the two terms exhibit identical rankings in two genres—namely, the TV/movie and academic genres. However, they display different rankings in the remaining six genres. This observation suggests that *show up* and *turn up* share a relatively modest similarity of approximately 25%, indicating a moderate level of differentiation between the two expressions in American English. Conversely, the ranking analysis in the BNC reveals that while *show up* and *turn up* share the same rank in the fiction genre, their rankings diverge in the other six genres, yielding a similarity rate of 14.28%. This discrepancy suggests a subtle yet significant national variation in the usage of these expressions between American and British English.

Additionally, it is important to note that in the COCA, *show up* is furthest from *turn up* in the fiction genre, whereas the two terms exhibit the closest proximity in the TV/movie genre. In contrast, in the BNC, the two expressions are most distantly ranked in the spoken genre, while they are closest in the magazine genre. These differences in genre-based rankings underscore notable national variations in how *show up* and *turn up* are employed in American and British English, further emphasizing the distinct linguistic patterns observed in each variety.

Furthermore, a closer examination of the frequencies of *show up* and *turn up* reveals distinct patterns in the two corpora. In the COCA, *show up* demonstrates a greater degree of variability, with frequencies deviating considerably from the mean. On the other hand, the frequencies of *turn up* in the COCA exhibit a more consistent pattern, with values tightly clustered around the average. In contrast, the BNC shows a different trend: *show up* frequencies exhibit a more stable and consistent usage pattern, while *turn up* frequencies fluctuate more significantly, displaying greater variability from the mean. These contrasting patterns of variability between American and British English further highlight the distinct usage characteristics of these two terms in each variety.

In terms of statistical analysis, the COCA yields a calculated correlation coefficient (r value) of 0.7375, indicating a strong positive correlation between the occurrences of *show up* and *turn up*. This suggests that as the frequency of one term increases, so does the frequency of the other, pointing to a clear and consistent relationship between their usage patterns in American English. In contrast, the BNC exhibits a significantly lower correlation coefficient of 0.0669, implying that there is no meaningful or consistent relationship between the frequencies of the two terms in British English. This low r value suggests that, unlike in American English, the usage of *show up* and *turn up* in British English is largely independent of each other.

It is also important to emphasize that, in the COCA, the frequency of *show up* plays a significant role in explaining the frequency of *turn up*, reflecting a clear, statistically significant relationship between the two terms in American English. By contrast, in the BNC, the frequency of *turn up* cannot be reliably predicted by the frequency of *show up*, highlighting a marked difference in the way these expressions are used across the two varieties of English.

In conclusion, while both the COCA and the BNC provide evidence of a relationship between *show up* and *turn up*, the strength and nature of this relationship vary notably between American and British English. In the COCA, the two terms exhibit a strong positive correlation, suggesting a significant association in their usage, while the BNC reveals a much weaker, almost negligible relationship. This variation points to important national differences in the way these expressions are utilized and understood, further underscoring the broader linguistic distinctions between American and British English.

6. Conclusions

This study examines the relationship between the phrasal verbs *show up* and *turn up* using data from the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). In COCA, the two expressions share the same ranking in two of eight genres—TV/movie and academic—yielding a similarity rate of 25%. In contrast, in the BNC, they match in only one of seven genres (fiction), resulting in a lower similarity of 14.28%. These findings suggest limited overlap in genre-based usage, pointing to notable differences between American and British English. Further analysis of genre proximity reveals that in COCA, *show up* and *turn up* differ most in the fiction genre and align most closely in the TV/movie genre. In the BNC, they are furthest apart in spoken discourse and most similar in magazines. These patterns underscore distinct usage tendencies across genres, reflecting broader national variation in English usage. Examining frequency distributions provides additional insight. In COCA, *show up* demonstrates greater variability, with frequencies spread widely across genres, while *turn up* shows a more stable and consistent pattern. In the BNC, the opposite is observed: *show up* remains relatively consistent, whereas *turn up* exhibits greater fluctuation. These contrasting trends further emphasize the differences in how each term functions within the two national varieties. Statistical analysis confirms these differences. In COCA, the correlation between the two expressions is strong ($r = 0.7375$), suggesting that their usage patterns tend to increase or decrease together across genres. However, in the BNC, the correlation is minimal ($r = 0.0669$), indicating that the two expressions are largely used independently in British English. Overall, these results highlight significant national variation in the use of *show up* and *turn up*. American English shows a stronger and more consistent relationship between the terms across genres, while British English reflects a weaker and more independent usage pattern. These differences—in rankings, frequency distributions, and statistical correlations—underscore how phrasal verbs can behave differently across varieties of English, reinforcing the importance of corpus-based analysis in understanding language variation.

Author Contributions

The first author, N.K., and the corresponding author, E.H.K., wrote this article together. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The overall frequency of *show up* and *turn up* was obtained from the COCA and the BNC, respectively. These data are still available on Google.

Acknowledgments

Our thanks are owed to Far East University in South Korea.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Aarts, B., Granger, S., 1993. Corpus-based approaches to grammar. Benjamins Publishing: Amsterdam, Netherlands.

[2] Baker, P., Hardie, A., McEnery, T., 2006. A glossary of corpus linguistics. Edinburgh University Press: Edinburgh, UK.

[3] Barlow, M., Kuperman, V., 2008. Corpus linguistics: A guide to the theory and practice. Palgrave Macmillan: New York, NY, USA.

[4] Biber, D., 1993. Representativeness in corpus design. Literary and Linguistic Computing. 8(4), 243–257. DOI: https://doi.org/10.1093/llc/8.4.243

[5] Biber, D., Conrad, S., Reppen, R., 1998. Corpus linguistics: Investigating language structure and use. Cambridge University Press: Cambridge, UK.

[6] British National Corpus (BNC), n.d. British National Corpus. Available from: https://corpus.byu.edu/bnc (cited 8 March 2025).

[7] Channell, J., 2000. Corpus linguistics: A resource book for students. Routledge: London, UK.

[8] Corpus of Contemporary American English (COCA), n.d. Corpus of Contemporary American English. Available from: https://corpus.byu.edu/coca (cited 8 March 2025).

[9] Corpus of Historical American English (COHA), n.d. Corpus of Historical American English. Available from: https://corpus.byu.edu/coha (cited 8 March 2025).

[10] Firth, J.R., 1957. Papers in linguistics 1934-1951. Oxford University Press: Oxford, UK.

[11] Gries, S.T., 2006. Exploring variability within and between corpora: Some methodological considerations. Corpora. 1(2), 109–151. DOI: https://doi.org/10.3366/cor.2006.1.2.109

[12] Gries, S.T., 2013. Statistics for linguists: A step-by-step guide. Mouton de Gruyter: Berlin, Germany.

[13] Hansard Corpus (HC), n.d. Hansard Corpus. Available from: https://english-corpora.org/hansard/ (cited 8 March 2025).

[14] Hunston, S., Francis, G., 2000. Pattern grammar: A corpus-driven approach to the lexical grammar of English. John Benjamins Publishing: Amsterdam, Netherlands.

[15] McEnery, T., Wilson, A., 2001. Corpus linguistics: An introduction. Edinburgh University Press: Edinburgh, UK.

[16] Meyer, C., 2002. English corpus linguistics: An introduction. Cambridge University Press: Cambridge, UK.

[17] O'Keeffe, A., McCarthy, M., Carter, R., 2007. From corpus to classroom: Language use and language teaching. Cambridge University Press: Cambridge, UK.

[18] Sinclair, J., 1991. Corpus, concordance, collocation. Oxford University Press: Oxford, UK.

[19] Tognini-Bonelli, E., 2001. Corpus linguistics at work. John Benjamins Publishing: Amsterdam, Netherlands.