BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

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

ARTICLE

Get on with and Continue with: Similarity Analysis

Namkil Kang , ***Hyewon Cho*** *

College of Liberal Arts, Far East University, 76-32 Daehak-gil, Gamgok-myeon, Eumseong-gun, Chungbuk 27737, Republic of Korea

ABSTRACT

This article provides an in-depth similarity analysis of the phrases *get on with* and *continue with* using data from the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), and ChatGPT. A key finding is that in COCA, the two phrases share a 33.33% similarity in ranking analysis, whereas in BNC, their similarity is 0%. In COCA, *get on with* is most similar to *continue with* in the newspaper genre and least similar in TV/movies. Conversely, in BNC, their closest similarity occurs in magazines and their greatest divergence in fiction. Standard deviation analysis further highlights differences in frequency. In COCA, *get on with* has a standard deviation of 294.02, indicating a frequency range of 199.48 to 787.52, while *continue with* has a standard deviation of 194.4, with a range of 163.98 to 552.78. Although their frequency correlation is not statistically significant, COCA shows a weak positive correlation, while BNC reveals a weak negative correlation. Notably, in neither corpus does the frequency of *get on with* significantly affect *continue with*. Additionally, eight of their top 20 collocations overlap, reflecting a 40% similarity in usage. Overall, the findings suggest minimal similarity between these phrases, with clear distinctions between American and British English. These insights contribute to a deeper understanding of how phrase usage varies across different linguistic and cultural contexts.

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

*CORRESPONDING AUTHOR:

Hyewon Cho, College of Liberal Arts, Far East University, 76-32 Daehak-gil, Gamgok-myeon, Eumseong-gun, Chungbuk 27737, Republic of Korea; Email: jypenn@hanmail.net

ARTICLE INFO

Received: 16 February 2025 | Revised: 15 March 2025 | Accepted: 21 March 2025 | Published Online: 30 March 2025
DOI: https://doi.org/10.30564/fls.v7i4.8764

CITATION

Kang, N., Cho, H., 2025. Get on with and Continue with: Similarity Analysis. Forum for Linguistic Studies. 7(4): 226–237. DOI: https://doi.org/10.30564/fls.v7i4.8764

1. Introduction

Our analyses reveal notable differences in the usage, frequency, and correlation of *get on with* and *continue with* across American and British English. While *get on with* and *continue with* share similar meanings, their actual usage patterns warrant closer examination, particularly across English varieties. This article aims to provide an in-depth similarity analysis of get on with and continue with using data from the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), and ChatGPT [1–3]. These expressions are selected due to their functional similarity in conveying the idea of maintaining or resuming an activity. However, despite their semantic proximity, they differ in formality, usage frequency, and potential cultural nuance. Understanding how these expressions are distributed and used across different contexts can provide insights into subtle linguistic preferences, pragmatic variation, and potential challenges for non-native speakers. To account for both regional and stylistic variation in English usage, this study employs two major corpora: BNC and COCA. The BNC offers a comprehensive representation of British English across a wide range of genres, while COCA provides a comparably balanced and genre-diverse portrayal of American English. The comparative use of these corpora allows for an exploration of how these expressions differ not only by context but also by regional dialect (British vs. American English), thereby enriching the linguistic analysis with both diachronic and sociolinguistic perspectives.

This research includes seven distinct analyses of *get on with* and *continue with*. The reason for conducting these seven different analyses is to explore the extent of similarity between the two expressions. Another goal of this article is to predict whether there is any national variation in the use of *get on with* and *continue with*. Note that the COCA is a representative corpus of American English, while the BNC is a representative corpus of British English. Major corpora such as the COCA, the BNC, the Hansard Corpus (HC), the Corpus of Historical American English (COHA), and studies by Murphy (2016, 2019) have provided significant linguistic insights [1, 2, 4–7]. Corpora help us better understand the subtle differences between *get on with* and *continue with*. We obtained the frequencies of *get on with* and *continue with* from the COCA and the BNC, which were used for the seven different analyses. For more detailed insights into corpus linguistics, readers may refer to some seminal works [8–28].

The study employs seven analytical approaches: (1) ranking analysis to explore how similar *get on with* and *continue with* are in terms of their rankings, (2) Euclidean distance measurement to assess whether the two expressions exhibit high similarity in each genre, (3) variance analysis to measure how much the individual data deviate from the mean, (4) standard deviation analysis to examine how much the relevant data vary from the mean,(5) correlation analysis to measure the degree of correlation between the frequencies of *get on with* and *continu with*, (6) linear regression analysis to observe whether the frequency of the independent variable, *get on with*, affects the frequency of the dependent variable, *continue with*, (7) collocation analysis to identify the collocations of *get on with* and *continue with*.

2. Materials and Methods

The main goal of this article is to provide seven in-depth analyses of *get on with* and *continue with*. We obtained the relevant data from the COCA, the BNC, and through ChatGPT. As mentioned earlier, we conducted seven analyses to examine how similar *get on with* and *continue with* are in the COCA and the BNC. More specifically, we conducted ranking analysis, Euclidean distance analysis, variance analysis, standard deviation analysis, correlation analysis, linear regression analysis, and collocation analysis, which allowed us to assess how closely *get on with* and *continue with* are related. This research conducts seven distinct analyses to explore the degree of similarity between *get on with* and *continue with* and to determine whether there is national variation in their usage. These seven analyses are as follows:

(1) Ranking Analysis

Evaluates how similar *get on with* and *continue with* are in terms of their rankings within the COCA and the BNC. By comparing their rankings across different genres, we gain insight into how closely the two expressions are related.

(2) Euclidean Distance Analysis

Measures the similarity between the two expressions across various genres. A smaller Euclidean distance suggests a higher degree of similarity in usage patterns.

(3) Variance Analysis

Examines the extent to which the frequency of each expression deviates from the mean. A higher variance indicates

greater fluctuation in usage across different genres.

(4) Standard Deviation Analysis

Assesses the degree of dispersion in the frequency values of *get on with* and *continue with*. Helps determine the consistency or variability of their usage within the COCA and the BNC.

(5) Correlation Analysis

Determines the relationship between the frequencies of *get on with* and *continue with*. A positive correlation suggests that as one expression's frequency increases, the other tends to increase as well, whereas a negative correlation indicates an inverse relationship.

(6) Linear Regression Analysis

Examines whether the frequency of *get on with* (independent variable) influences the frequency of *continue with* (dependent variable). A stronger positive or negative regression coefficient indicates a stronger relationship, while values closer to zero suggest a weaker association.

(7) Collocation Analysis

Identifies the most common collocations of *get on with* and *continue with*. ChatGPT provides the top 20 collocations in descending order, which are analyzed to determine their contextual similarities.

Together, these analyses provide a comprehensive comparison of *get on with* and *continue with*, offering valuable insights into their linguistic behavior and potential national variation between American and British English.

3. Data Collection

We obtained the relevant data through COCA and BNC and ChatGPT. The COCA consists of eight primary genres:

(1) Spoken– Transcriptions of unscripted conversations from TV, radio, and other spoken sources.
(2) Fiction– Novels, short stories, and scripts from books, magazines, and TV/movie scripts.
(3) Magazine– Articles from a variety of popular magazines.
(4) Newspaper– Articles from major American newspapers covering different topics.
(5) Academic– Scholarly and research-based articles from academic journals.
(6) TV/Movies– Dialogues and scripts from television shows and films.
(7) Blog–Online texts from blogs.
(8) Web– Online texts from websites.

On the other hand, the BNC contains seven primary genres:

(1) Spoken– Informal conversations, interviews, and transcriptions of spontaneous speech.
(2) Fiction– Books, short stories, and scripts from British literature.
(3) Magazine– Articles from British periodicals and magazines.
(4) Newspaper– News articles from various UK newspapers.
(5) Academic– Scholarly texts, including research papers and textbooks.
(6) Non-Academic (Informative)– Non-fiction books and manuals (e.g., biographies, travel guides).
(7) Miscellaneous– Other written texts that do not fall into the main categories.

The key difference between the COCA and the BNC is that the COCA has a dedicated TV/Movies category, while the BNC includes Non-Academic Informative as a separate category. The COCA also includes Blog/Web data, reflecting modern digital usage, which is absent in the BNC. Additionally, we obtained the top 20 collocations of *get on with* and *continue with* through ChatGPT. It is worth noting that ChatGPT provided them in descending order, which was used for collocation analysis.

4. Results

4.1. Ranking Analysis

This section is dedicated to investigating the difference between *get on with* and *continue with* in the COCA and the BNC, which can enables us to figure out the degree of the similarity between them. Our ranking analysis refers to the overall frequency order of each genre. This analysis makes it easier to see how different *get on with* and *continue with* are in their rankings. Now have a look at **Table 1**.

Perhaps it is worthwhile pointing out that *get on with* and *continue with* have the highest frequency and proportion in the TV/movie genre and the spoken genre, respectively. This, in turn, suggests that *get on with* was the most widely used in the TV/movie genre, while *continue with* was the most commonly used in the spoken genre. Quite interestingly,

Table 1. Ranking Analysis in the COCA.

Ranking	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8
Get on with (frequency)	TV/M (1,065)	Fic	Spok (660)	Web (413)	Blog (383)	News (310)	Mag (278)	Acad (83)
Continue with (frequency)	Spok (776)	Blog (480)	Web (424)	TV/M (294)	Acad (291)	News (238)	Mag (187)	Fic

Note. TV/M= TV/Movie, Fic = Fiction, Spok = Spoken, Mag = Magazine, Acad = Academic.

Americans prefer to use *continue with* in daily conversation. It must be noted that *continue with* is preferred over *get on with* in the spoken genre. As exemplified in Table 1, the TV/movie genre was the most influenced by *get on with*, closely followed by the fiction genre, the spoken genre, the web genre, the blog genre, the newspaper genre, the magazine genre, and the academic genre, in that order. On the other hand, the spoken genre was the most influenced by *continue with*, closely followed by the blog genre, the web genre, the TV/movie genre, the academic genre, the newspaper genre, the magazine genre, and the fiction genre, in descending order. This, in turn, indicates that *get on with* and *continue with* reveal the same rankings in only the newspaper genre and the magazine genre, while they exhibit different rankings in the other six genres. From all of this, it seems clear that *get on with* and *continue with* are 33.33% similar in their ranking analysis. Put differently, they exhibit a low degree of similarity in their ranking analysis.

Now let us turn to the rankings of *get on with* and *continue with* in seven genres of the BNC (**Table 2**).

It is probably worthwhile pointing out that *get on with* and *continue with* have the highest frequency and proportion in the fiction genre and the miscellaneous genre, respectively. Quite interestingly, *get on with* was the most frequently used in the fiction genre, but *continue with* was not. As illustrated in **Table 2**, the fiction genre was the most influenced by *get on with*, closely followed by the spoken genre, the miscellaneous genre, the newspaper genre, the non-academic genre, the magazine genre, and the academic genre, in that order. The miscellaneous genre was the most influenced by *continue with*, closely followed by the non-academic genre, the newspaper genre, the academic genre, the spoken genre, the fiction genre, and the magazine genre. More interestingly, *get on with* and *continue with* reveal different rankings in all the genres, which, in turn, indicates that they are 0% similar in their ranking analysis. It therefore seems reasonable to conclude that *get on with* and *continue with* reveal a low similarity in American English, whereas they exhibit no similarity in British English with respect to their rankings. Simply put, there seems to be national variation in the use of *get on with* and *continue with*.

The ratio values in **Table 3** represent the relative frequency ranking of *get on with* compared to *continue with* in each genre. A higher ratio (>1) indicates that *get on with* is used more frequently than *continue with*, while a lower ratio (<1) suggests that *continue with* is more dominant. By comparing the values across genres in COCA and BNC, we can observe distinct patterns in usage. In COCA, narrative-based genres such as Fiction (4.27) and TV/Movie (3.621) show a strong preference for *get on with* over *continue with*. However, *continue with* is used much more frequently than *get on with* in academic writing. Spoken language shows a slightly higher usage of *continue with* than *get on with*. In BNC, Fiction (10.622) and Spoken (6.19) show an extremely high preference for *get on with* over *continue with*. Only in Academic texts does *continue with* appear more frequently than *get on with*. The higher ratio of *get on with* in BNC suggests that it is a more common expression in British English, whereas American English tends to use *continue with* more frequently in certain contexts. The differences highlight regional preferences in phrase usage across genres.

4.2. Euclidean Distance Analysis

In what follows, we aim to explore the actual distance between *get on with* and *continue with* in eight genres and seven genres. We will employ Euclidean distance to investigate the distance between *get on with* and *continue with* in eight genres and seven genres. Instead of using raw frequency counts, we normalize the data by calculating the percentage of occurrences of each phrase within each genre relative to its total occurrences in the corpus. Note that closer the distance between *get on with* and *continue with*, the higher the degree of similarity. We define Euclidean distance in:

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}$$

Table 2. Ranking Analysis in the BNC.

Ranking	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7
Get on with (frequency)	Fic (393)	Spok (260)	Misc (205)	News (146)	Non-acad (89)	Mag (71)	Acad (55)
Continue with (frequency)	Misc (147)	Non-acad (74)	News (66)	Acad (48)	Spok (42)	Fic (37)	Mag (31)

Note. Fic = Fiction, Spok = Spoken, Non-acad = Non-academic, Mag = Magazine, Acad = Academic,Misc = Miscellaneous.

Table 3. Ranking of *Get on with* compared to *Continue with* based on Ratio Values.

	Genre	*Get on with*	*Continue with*	Ratio
COCA	Fiction	756	177	4.27
	TV/Movie	1065	294	3.621
	Magazine	278	187	1.485
	Newspaper	310	238	1.302
	Web	413	424	0.974
	Spoken	660	776	0.851
	Blog	383	480	0.798
	Academic	83	291	0.285
BNC	Fiction	393	37	10.622
	Spoken	260	42	6.19
	Newspaper	146	74	1.973
	Non-academic	89	48	1.854
	Miscellaneous	205	147	1.395
	Magazine	71	66	1.077
	Academic	55	74	0.743

where p_i represents the percentage of *get on with* in genre *i.* q_i represents the percentage of *continue with* in genre *i. n* is the number of genres.

Now have a look at **Table 4**:

Table 4. Euclidean Distance in the COCA.

Genre	Spok	Fic	Mag	News	Acad	TV/M	Blog	Web	TOTAL
Get on with (%)	16.72	19.15	7.04	7.85	2.10	26.98	9.70	10.46	100
Continue with (%)	27.07	6.17	6.52	8.30	10.15	10.25	16.74	14.79	100
Euclidean Distance	10.35	12.97	0.52	0.45	8.04	16.72	7.04	4.25	

It is interesting to point out that *get on with* is closest to *continue with* in the newspaper genre. More specifically, the distance between *get on with* and *continue with* is 0.45, thus indicating that they have the highest similarity in the newspaper genre. It is worthwhile to consider the magazine genre. It is worth noticing that the newspaper genre is followed by the magazine genre. *Get on with* is the second closest to *continue with* in the magazine genre. The distance between them is 0.52, which, in turn, implies that they are the second highest similarity in the magazine genre. Particularly noteworthy is the fact that *get on with* is furthest from *continue with* in the TV/movie genre. To be more specific, the distance between them is 16.72, hence implying that they exhibit the lowest similarity. It is worthwhile to include the fiction genre. Quite interestingly, *get on with* is the second furthest to *continue with* in the fiction genre. This seems to suggest that they show the second lowest similarity in the fiction genre. We thus conclude that *get on with* is closest to *continue with* in the newspaper genre, while the former is furthest from the latter in the TV/movie genre. Put differently, they exhibit the highest similarity in the newspaper genre, while they reveal the lowest similarity in the TV/movie genre.

Now let us turn our attention to the BNC (**Table 5**):

In **Table 5**, it is worth mentioning that *get on with* is closest to *continue with* in the magazine genre. To be more specific, the actual distance between them is 1.14, thus exhibiting the highest similarity. What is interesting is that the magazine genre is followed by the newspaper genre. More

Table 5. Euclidean Distance in the BNC.

Genre	Spok	Fict	Mag	News	Acad	Non-Acad	Misc	TOTAL
Get on with (%)	21.33	32.24	5.82	11.98	4.51	7.30	16.82	100
Continue with (%)	9.44	8.31	6.97	14.83	10.79	16.63	33.03	100
Euclidean Distance	11.89	23.92	1.14	2.86	6.27	9.32	16.22	

specifically, in the newspaper genre, *get on with* is the second closest to *continue with*, thus indicating that they exhibit the second highest similarity. More interestingly, *get on with* is furthest from *continue with* in the fiction genre. The actual distance between them is 23.92, which, in turn, implies that they reveal the lowest similarity in the fiction genre. It must be emphasized, on the other hand, that *get on with* is the second furthest from *continue with* in the miscellaneous genre (16.22), thus suggesting that they exhibit the second lowest similarity. It therefore seems clear that there is national variation in the use of *get on with* and *continue with*. More specifically, *get on with* is closest to *continue with* in the newspaper genre in the case of the COCA, whereas the former is furthest from the latter in the TV/movie genre. On the other hand, *get on with* is closest to *continue with* in the magazine genre in the case of the BNC, whereas the former is furthest from the latter in the fiction genre. Simply put, there seems to be national variation between American English and British English.

4.3. Variance Analysis and Standard Deviation Analysis

In what follows, we aim to account for the frequencies of *get on with* and *continue with* in terms of the variance and standard deviation. The variance accounts for the overall spread of data and helps understand the degree of variability or consistency within a set of frequencies. On the other hand, the standard deviation measures the amount of variation or dispersion in a set of data. Specifically, it quantifies how much individual data deviate from the mean (average). Let us have a look at **Table 6**:

Table 6. Frequency of *Get on with* and *Continue with* in the COCA.

Get on with (frequency)	TV/M (1,065)	Fic (756)	Spok (660)	Web (413)	Blog (383)	News (310)	Mag (278)	Acad (83)
Continue with (frequency)	Spok (776)	Blog (480)	Web (424)	TV/M (294)	Acad (291)	News (238)	Mag (187)	Fic (177)

Also, let us have a look at **Table 7**:

Table 7. The Mean, Variance, and Standard Deviation of *Get on with* and *Continue with* in the COCA.

Type	Mean	Variance	Standard Deviation
Get on with	493.5	86,232.25	294.02
Continue with	358.375	37,717.91	194.4

As exemplified in **Table 7**, the variance of the frequencies of get on with is 86,232.25, indicating a high degree of spread around the mean. The standard deviation of approximately 294.02 further confirms that the values vary significantly from the mean. The standard deviation of the frequencies of get on with is 294.02, which, in turn, indicates that the frequencies of get on with roughly fall within the range of approximately 493.5 ± 294.02 . On the other hand, the standard deviation of the frequencies of continue with is 194.4, which, in turn, suggests that the frequencies of continue with roughly fall within the range of approximately 358.375 ± 194.4. There is a large spread between the frequencies in the case of get on with, indicating that some frequencies (such as 1,065 and 278) are much higher or lower than the mean. On the other hand, the average frequency in the case of continue with is 358.375, indicating the central

tendency of the data. The variance of 37.717.91 reflects a moderate level of dispersion around the mean. This suggests that while the values are spread out, they are not excessively varied. Now take a look at **Table 8**:

Table 8. Frequency of *Get on with* and *Continue with* in the BNC.

Get on with (frequency)	Fic (393)	Spok (260)	Misc (205)	News (146)	Non-acad (89)	Mag (71)	Acad (55)
Continue with (frequency)	Misc (147)	Non-acad (74)	News (66)	Acad (48)	Spok (42)	Fic (37)	Mag (31)

Also, let us take a look at **Table 9**:

Table 9. The Mean, Variance, and Standard Deviation of *Get on with* and *Continue with* in the BNC.

Type	Mean	Variance	Standard Deviation
Get on with	174.14	12,589.39	112.19
Continue with	63.57	1,358.39	36.85

As indicated in **Table 9**, the average frequency in the case of *get on with* is 174.14. The variance of the frequencies of *get on with* is 12,589.39, which indicates the level of dispersion in the data. This suggests a moderate spread of values around the mean. The standard deviation of approximately 112.19 suggests that the values deviate by around 112 from the mean. More specifically, the standard deviation of the frequencies of *get on with* is 112.19, which, in turn, indicates that the frequencies of *get on with* roughly lie between 174.14 – 112.19 and 174.14 + 112.19. Quite interestingly, some values (such as 393) are much higher than the mean, while others (such as 55) are much lower. This suggests that there are a few extreme values influencing the overall variance. The data shows moderate to high variability, with frequencies ranging from as low as 55 to as high as 393. The calculated variance and standard deviation indicate that while the values are spread out around the mean, there is no extremely outliers compared to the mean of the data. In the case of *continue with*, the variance of its frequencies is relatively high. A value of 36.85 suggests that most data lie within about 36.85 units of the mean. More specifically, the standard deviation of the frequencies of *continue with* is 36.85, which, in turn, implies that the frequencies of continue with roughly fall within the range of approximately 63.57 ± 36.85. To sum up, the use of *get on with* in America seems to be different from that of *get on with* in the UK in that the variance and standard deviation of *get on with* in the COCA are extremely high, but those of *get on with* in the BNC are relatively high. A higher standard deviation indicates that data are spread out more widely from the mean, while a lower standard deviation means that data are closer to the mean. If a dataset has a small standard deviation, the values are more consistent and less variable. Conversely, a large standard deviation indicates more variability in the data. This, in turn, indicates that in America, there is large variation in the use of *get on with* across different genres, whereas such variation is not observed in the UK.

4.4. Correlation Analysis

This section is devoted to providing correlation analysis through which we can grasp whether or not the frequencies of *get on with* and those of *continue with* are correlated. We already computed the Pearson correlation coefficient (*r*) between the frequencies of *get on with* and *continue with* in the COCA which was approximately 0.121 (**Table 10**).

Table 10. Pearson Correlation Coefficient (r).

		continue with
get on with	*r*	0.121
	p	0.775
	N	8

The Pearson correlation coefficient measures the strength and direction of the linear relationship between two variables. The value of *r* ranges from −1 to +1. The strength of the correlation is typically categorized as follows

(**Table 11**).

Table 11. 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

In our case, the correlation coefficient *r* is 0.121and it is close to 0, indicating a very weak positive correlation (see **Table 12**). This suggests that, although there is some positive relationship between the two set of frequencies, the strength of the relationship is weak. In other words, the frequencies of *get on with* and *continue with* are not strongly related to each other. However, it was not statistically significant (p = 0.775).

Table 12. Steps for Significance Testing.

Null Hypothesis	There is no linear relationship between the two variables (i.e., r = 0)
Alternative Hypothesis	There is a linear relationship between the two variables (i.e., r ≠ 0)

The p-value is used to test the null hypothesis. If the p-value is less than 0.05, we reject the null hypothesis, meaning the correlation is statistically significant. If it is greater than 0.05, we fail to reject the null hypothesis, meaning the correlation is not statistically significant. Note that the p-value for the correlation between the frequencies of *get on with* and *continue with* is approximately 0.775. A p-value of 0.775 is much higher than 0.05. This means that we fail to reject the null hypothesis, indicating that the correlation is not statistically significant.

Now let us turn our attention to the correlation of the frequencies of *get on with* and *continue with* in the BNC. We computed the Pearson correlation coefficient (*r*) between the frequencies of *get on with* and *continue with*, which was approximately −0.029 (**Table 13**).

Table 13. Pearson Correlation Coefficient (r).

		continue with
get on with	*r*	−0.029
	P	0.951
	N	7

Notice that the Pearson correlation coefficient is approximately −0.029 in the BNC. This indicates a weak negative correlation between the frequencies of *get on with* and *continue with*. In other words, as the frequency of one phrase increases, the frequency of the other phrase tends to decrease slightly, but the relationship is not strong. Note that the p-value for the correlation between *get on with* and *continue with* is approximately 0.951. Since the p-value is greater than 0.05, we fail to reject the null hypothesis, meaning the correlation is not statistically significant. Again, there is national variation between the frequencies of *get on with* and *continue with*. Even though statistically significant relationships were not found, it seems that there might be a weak positive correlation between the frequencies of *get on with* and *continue with* in the COCA, while there might be a weak negative correlation between them in the BNC. More specifically, as the frequency of one phrase increases, the frequency of the other phrase tends to increase slightly in the COCA, while, as the frequency of one phrase increases, the frequency of the other phrase tends to decrease slightly in the BNC.

4.5. Linear Regression Analysis

In what follows, we aim to provide linear regression analysis through which we can observe whether the frequency of the independent variable *get on with* affects the frequency of the dependent variable *continue with*. Note that *get on with* is an independent variable, while *continue with* is a dependent variable. The so-called covariance measures the relationship between two variables and tells us whether they tend to increase or decrease together. The covariance between *get on with and continue with* is approximately 5893.58. This seems to say that in the case of the COCA, the two variables tend to increase together. For more details, let us have a look at **Tables 14** and **15**:

The regression analysis with the frequency of *get on with* as the independent variable and the frequency of *continue with* as the dependent variable in the COCA corpus showed that the model explained 1.5% of the variance (R^2 = 0.015), indicating poor model fit. The ANOVA results

Table 14. Model Summary of COCA.

Model	R	R^2	Adjusted R^2	Std. Error
1	0.121	0.015	−0.150	214.036

Table 15. Coefficients.

Model	B	Std. Error	Beta	*t*	*p*
1 (Constant)	320.436	147.929		2.166	0.073
get on with	0.077	0.258	0.121	0.298	0.775

revealed that the overall model was not statistically significant (p = 0.775). Furthermore, the regression coefficient was B = 0.077, but it was not statistically significant (p = 0.775), suggesting that the frequency of *get on with* had no significant effect on the frequency of *continue with*.

Now let us turn our attention to the BNC. Let us have a look at **Tables 16** and **17**:

Table 16. Model Summary.

Model	R	R^2	Adjusted R^2	Std. Error
1	0.029	0.001	−0.199	43.685

Table 17. Coefficients.

Model	B	Std. Error	Beta	*t*	*p*
1 (Constant)	65.203	30.367		2.147	0.085
get on with	−0.009	0.146	−0.029	−0.064	0.951

The regression analysis on the frequency of *get on with* and *continue with* in the BNC corpus showed that the model explained only 0.1% of the variance ($R^2 = 0.001$), indicating a very poor model fit. The ANOVA results revealed that the overall model was not statistically significant (p = 0.951). Additionally, the regression coefficient was B = −0.009, with a p-value of 0.951, suggesting that the frequency of *get on with* did not have a significant impact on that of *continue with*.

4.6. Collocation Analysis

The goal of this section is to probe into the collocations of *get on with* and *continue with*. We obtained the top 20 collocations of *get on with* and *continue with* through ChatGPT. Let us take a look at **Table 18**:

It is important to note that the word *work* is the first collocation suggested by ChatGPT. Specifically, it is the most frequently used word with both *get on with* and *continue with*. This, in turn, indicates that *work* was the most commonly occurring term in the web data from which ChatGPT learned. Notably, the words *life* and *project* rank second in frequency, being the most common collocates of *get on with* and *continue with*, respectively. Furthermore, the words *task* and *plan* are the third most frequently suggested collocates for *get on with* and *continue with*. This suggests that these terms were the third most prevalent in web usage alongside these two phrases. It is also noteworthy that *project* and *meeting* are the fourth most suggested collocates, indicating their prominence in web data in relation to these expressions. Additionally, *job* and *discussion* appear as the fifth most common collocates, further supporting the notion that these words were frequently used in conjunction with *get on with* and *continue with* in the data. To summarize, *work* is the most commonly used word with *get on with*, followed by *life*, *task*, *project*, and *job*. For *continue with*, the most frequent collocates are *work*, *project*, *plan*, *meeting*, and *discussion*, in descending order of frequency. Most importantly, 8 out of the top 20 collocations of *get on with* and *continue with*

Table 18. Collocations of Get on with and Continue with.

Number	Collocations of *Get on with*	Collocations of *Continue with*
1	get on with work	continue with work
2	get on with life	continue with the project
3	get on with the task	continue with the plan
4	get on with the project	continue with the meeting
5	get on with the job	continue with the discussion
6	get on with studying	continue with the task
7	get on with your day	continue with business
8	get on with business	continue with the process
9	get on with the program	continue with the research
10	get on with the work	continue with the study
11	get on with the discussion	continue with the program
12	get on with research	continue with the development
13	get on with the plan	continue with the review
14	get on with the meeting	continue with the repairs
15	get on with homework	continue with the project
16	get on with the repair	continue with the investigation
17	get on with the presentation	continue with the negotiations
18	get on with the decision	continue with the training
19	get on with the preparations	continue with the presentation
20	get on with the event	continue with the decision

are identical, suggesting that these two phrases share 40% of their top 20 collocations. This finding points to a relatively low degree of similarity in the collocational patterns of *get on with* and *continue with*.

5. Discussion

In this analysis, we examine key findings from seven distinct studies on the phrases *get on with* and *continue with*. These insights provide a comparative overview of their usage across different genres, highlighting both similarities and differences.

First, an analysis of the COCA shows that *get on with* and *continue with* share the same rankings only in the newspaper and magazine genres, while their rankings differ across the other six genres. This translates to a 33.33% similarity in ranking patterns, indicating a relatively low degree of alignment in their usage. In contrast, in the BNC, the two phrases have distinct rankings in all genres, yielding a 0% similarity. These findings suggest that while *get on with* and *continue with* exhibit limited similarity in American English, they show no alignment in British English, pointing to clear national variation in usage.

Second, the proximity between *get on with* and *continue with* differs across genres. In COCA, they are most similar in the newspaper genre and most divergent in the TV/movie genre. In the BNC, they are closest in the magazine genre and most distinct in fiction. These differences further support the notion of national variation in contextual usage.

Third, the use of *get on with* varies more in American English than in British English, as indicated by differences in variance and standard deviation across COCA and BNC. In COCA, *get on with* has an extremely high standard deviation, suggesting greater variability across genres, whereas in BNC, the standard deviation is high but comparatively lower. This indicates that the phrase is used more consistently in British English than in American English.

Fourth, correlation analysis reveals weak and statistically insignificant relationships between the frequencies of *get on with* and *continue with* in both corpora. In COCA, the Pearson correlation coefficient is 0.121 ($p = 0.775$), indicating a very weak positive correlation. In BNC, the coefficient is −0.029 ($p = 0.951$), showing a weak negative correlation. While not statistically significant, the trends suggest that as the frequency of one phrase increases, the other shows slight movement in opposite directions across different regional varieties of English.

Fifth, regression analysis confirms that *get on with* does not significantly predict *continue with* in either corpus. In

COCA, the model explains only 1.5% of the variance (R^2 = 0.015) and is not statistically significant (p = 0.775), with a regression coefficient (B = 0.077). In BNC, the model explains just 0.1% of the variance (R^2 = 0.001) and is also not statistically significant (p = 0.951), with a regression coefficient (B = −0.009). These findings indicate no meaningful relationship between the two phrases in either corpus.

Sixth, an analysis of common collocates reveals partial overlap between *get on with* and *continue with*. The most frequent collocate of *get on with* is *work*, followed by *life*, *task*, *project*, and *job*. For *continue with*, the most frequent collocates are *work*, *project*, *plan*, *meeting*, and *discussion*. Eight out of the top 20 collocates are shared, indicating a 40% similarity. While this suggests some lexical overlap, the overall collocational patterns remain relatively distinct.

In conclusion, *get on with* and *continue with* exhibit a low degree of similarity across ranking analysis, variance and standard deviation measures, collocation analysis, and correlation analysis, with no significant relationship identified through regression analysis. The findings also highlight notable national variations in six of the seven analyses, with the exception of the regression results. Overall, the relationship between *get on with* and *continue with* is weak, statistically insignificant, and subject to regional differences between American and British English.

6. Conclusions

In summary, we conducted seven detailed analyses of *get on with* and *continue with* using data from the COCA, the BNC, and ChatGPT. These analyses reveal notable differences in the usage, frequency, and correlation of the two phrases across American and British English. A key finding is that in COCA, *get on with* and *continue with* exhibit a 33.33% similarity in ranking analysis, whereas in BNC, they show no similarity, with a 0% overlap. This suggests that while the phrases may be somewhat interchangeable in American English, they are distinctly used in British English. Additionally, their genre-specific usage varies. In COCA, *get on with* is closest to *continue with* in the newspaper genre and most distant in the TV/movie genre. In contrast, in BNC, they are most similar in magazines but furthest apart in fiction. Further analysis of frequency distributions reveals notable differences. In COCA, the standard deviation for *get on with* is 294.02, with a frequency range of 199.48 to 787.52, while *continue with* has a standard deviation of 194.4, ranging from 163.98 to 552.75. In BNC, *get on with* shows a standard deviation of 112.19, suggesting a frequency range between 61.95 and 286.33. These variations highlight the differing prominence of these phrases in each corpus. Additionally, correlation analysis suggests weak relationships between the phrases. In COCA, a weak positive correlation is present, meaning that as the frequency of one phrase increases, the other tends to rise slightly as well. Conversely, in BNC, the correlation is weakly negative, indicating that an increase in one phrase's frequency corresponds with a slight decrease in the other's usage. Notably, in both corpora, *get on with* does not significantly influence the frequency of *continue with*, reinforcing their distinct usage patterns. Finally, collocation analysis provides further insight into their similarities and differences. Eight out of the top 20 collocations for *get on with* and *continue with* are identical, reflecting a 40% similarity in their most common word associations. While this indicates some overlap in meaning, the limited shared collocations further support the conclusion that these phrases are not direct substitutes and are used differently in American and British English. Overall, these findings underscore the nuanced distinctions between *get on with* and *continue with*, emphasizing their varied usage across linguistic and cultural contexts.

Author Contributions

N.K., and H.C., wrote this article together. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The overall frequency of *get on with* and *continue with* was obtained from the COCA and the BNC, respectively. These are still available on Google.

Acknowledgements

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

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Corpus of Contemporary American English (COCA). Available from: https://corpus.byu.edu/coca (cited 17 January 2025).

[2] British National Corpus (BNC), 2025. Available from: https://corpus.byu.edu/bnc (cited 17 January 2025).

[3] ChatGPT. Available from: https://chat.openai.com (cited 17 January 2025).

[4] Hansard Corpus (HC). Available from: https://english-corpora.org/hansard/ (cited 17 January 2025).

[5] Corpus of Historical American English (COHA). Available from: https://corpus.byu.edu/coha (cited 17 January 2025).

[6] Murphy, R., 2016. Grammar in Use. Cambridge University Press: Cambridge, UK.

[7] Murphy, R., 2019. English Grammar in Use. Cambridge University Press: New York, NY, USA.

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

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

[10] 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

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

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

[13] Casal, J.E., Shirai, Y., Lu, X., 2022. English verb-argument construction profiles in a specialized academic corpus: Variation by genre and discipline. English for Specific Purposes. 66, 94–107.

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

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

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

[17] Kilgarriff, A., Grefenstette, G., 2003. Introduction to the special issue on corpus-based approaches to lexical semantics. Computational Linguistics. 29(4), 493–500. DOI: https://doi.org/10.1162/089120103322753347.

[18] Kovaliuk, Y., 2023. Idioms in general English corpora: on frequency, register, and cross-variety variation. Respectus Philologicus. 44, 11–24.

[19] Leech, G., Fallon, P., 1992. Computer corpora and linguistic theory. Language. 68(3), 459–482. DOI: https://doi.org/10.2307/415793.

[20] Marušić, B., 2023. Multi-word expressions in annual reports of American and British corporations: a corpus-based diachronic study. Journal of Teaching English for Specific and Academic Purposes. 793–811.

[21] Matsumoto, N., 2020. A genre-based analysis of evaluative modality in multi-verb sequences in English. In: Hohaus, P., Schulze, R. (eds.). Re-Assessing Modalising Expressions. John Benjamins Publishing Company: Amsterdam, The Netherlands. pp. 225–252.

[22] McEnery, T., Kawai, Y., 2008. Collocational patterning in English: A corpus-based study. Language and Computers. 63, 75–88.

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

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

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

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

[27] Tognini-Bonelli, R., 1993. Interpretative Nodes in Discourse — Actual and Actually. In: Baker, M., Francis, G., Tognini-Bonelli, E. (eds.). Text and Technology: In honour of John Sinclair. John Benjamins Publishing Company: Amsterdam, Netherlands. 193–212. DOI: https://doi.org/10.1075/z.64.13tog

[28] Wasserman, L., 2004. All of Statistics: A Concise Course in Statistical Inference. Springer: Berlin, Germany. DOI: https://doi.org/10.1007/978-0-387-21736-9