

A Corpus-Based Analysis of the Descriptive verbs “Describe” and “Illustrate” in BAWE and BASE Corpora

Adinda Maghfirah

adindamaghfirah@students.undip.ac.id

Universitas Diponegoro

Prihantoro

prihantoro@lecturer.undip.ac.id

Universitas Diponegoro

Abstract

Descriptive verbs such as *describe* and *illustrate* are regularly used in academic English to express meanings, definition, and arguments. Yet their nuanced usage across spoken and written contexts often creates challenges for students. This study investigates the grammatical relations and collocational patterns of *describe* and *illustrate* in the British Academic Written English (BAWE) and British Academic Spoken English (BASE) corpora and how its function differently in academic contexts. This study utilizes Sketch Engines Word Sketch Difference and N-Gram tools, the research systematically analyzes noun and adverb collocates as subjects, objects, and modifiers, as well as multi-word phrase patterns. Results indicate that in BAWE, *describe* is predominantly associated with concrete noun subjects, whereas *illustrate* collocates more with abstract nouns in BASE. This research fills a gap by directly contrasting spoken and written academic registers, providing a phrase bank that can help student both in spoken and written text.

Keywords: Describe; illustrate; academic English; BAWE; BASE; spoken vs. written English.

Introduction

There are a number of descriptive verbs in English, such as *describe* and *illustrate*, which play important roles in academic contexts. These verbs are frequently used to present explanations, provide definitions, and support arguments in both spoken and written communication. However, although *describe* and *illustrate* are often considered synonymous, students may find it challenging to choose the appropriate verb based on context. Referring to the definitions from the Longman Dictionary of Contemporary English and the Merriam-Webster Dictionary, the verbs appear interchangeable. However, Liu (2023) argues that students often rely on dictionary-based synonym definitions, which may not accurately reflect real-world usage. Therefore, corpus-based research offers an empirical approach to understanding how synonymous descriptive verbs like *describe* and *illustrate* actually function in academic discourse.

Describe and *illustrate* are commonly used in academic contexts whether in writing essays, papers, reports, or during presentations. Despite their importance, students often struggle to choose the appropriate verb, especially since many academic verbs are semantically related but used differently depending on the context. Previous studies have mainly focused on general frequencies and meanings but have rarely explored how these

verbs differ in usage across academic spoken and written texts. This research is therefore significant, as it compares the use of both verbs using authentic academic data from BASE and BAWE corpora. The findings aim to help students and educators better understand how to use these verbs accurately and naturally in academic English. While prior research often utilized large-scale corpora containing mixed genres, this study uses specialized academic corpora to provide a more focused and unbiased analysis. Furthermore, this study integrates tools from Sketch Engine, such as Word Sketch Difference, N-grams, and concordances, which enable a systematic and time-efficient analysis. These tools also offer additional insights that reduce the need for manual categorization and deepen the understanding of real academic usage.

Several recent corpus-based studies have explored how synonymous verbs behave across different genres, focusing on their frequency, collocational patterns, and grammatical roles (Chaengchenkit, 2023; Narkprom, 2024; Jarrar et al., 2024). In addition, research by Pratiwi (2020) and Orhon et al. (2018) shows that even similar words, such as *however* or verb pairs like *refuse* and *reject*, reveal distinct grammatical and contextual preferences depending on whether they appear in written or spoken forms. Similarly, Szczygłowska (2023) examined the use of specificity adjectives in the BAWE corpus, while Nesi and Gardner (2018) analyzed the generic structure and linguistic features of successful academic writing in the same corpus. However, many of the aforementioned studies rely on large, general-purpose corpora such as COCA (Chaengchenkit, 2023; Narkprom, 2024; Selmistraitis, 2023; Pratiwi, 2020; Liu, 2023) or BNC (Jarrar et al., 2024), with fewer focusing on discipline-specific corpora that contrast spoken and written academic English. To address this gap, the present study investigates the use of *describe* and *illustrate* in two complementary academic corpora: the British Academic Written English (BAWE) and the British Academic Spoken English (BASE). Hong (2019) emphasizes the importance of examining grammatical relations and collocational patterns of such verbs through phraseological studies.

Chon and Shin (2013) also noted a research limitation in their study and suggested further investigation using both BAWE and BASE to identify academic collocations more comprehensively. This research offers novelty by combining both spoken and written academic corpora (BASE and BAWE), unlike previous studies that rely on only one type. It also focuses specifically on the grammatical functions of nouns (as subjects or objects) and adverbs (as modifiers). Additionally, by employing the Word Sketch Difference and N-gram tools, this study provides a comparative analysis of usage patterns and identifies authentic academic phrases used in real contexts.

Methodology

This study adopts a qualitative corpus-based design to investigate the usage of the descriptive verbs *describe* and *illustrate* in academic English. The research focuses on analyzing how these verbs function grammatically and contextually across two academic corpora: the British Academic Written English (BAWE) and the British Academic Spoken English (BASE). By examining grammatical relations and multi-word expressions (MWEs) associated with the two verbs, the study aims to provide insight into their use in real academic contexts, both spoken and written.

As a corpus-based study, this research does not involve human participants or a physical site. Instead, it relies on corpora available via the Sketch Engine, a web-based corpus analysis tool widely used in linguistic research. The BAWE and BASE corpora serve as the primary data sources.

The data for this study were drawn from two academic corpora available in Sketch Engine. BAWE (British Academic Written English): A corpus of 2,761 proficient student writing samples produced at UK universities, ranging from 500 to 5,000 words in length, totaling approximately 6.9 million words. BASE (British Academic Spoken English): A spoken academic corpus comprising 160 lectures and 3 seminars recorded at the University of Warwick and the University of Reading between 2000–2005, totaling approximately 1.47 million words. Both corpora represent a range of academic disciplines, including arts and humanities, life sciences, physical sciences, and social sciences. These corpora were selected because they offer clear contrasts between written and spoken registers of academic English.

This study investigates two semantically related academic verbs: *describe* and *illustrate*. These verbs were selected due to their frequent use in academic genres, such as thesis, research papers, and presentations. Focusing on only two verbs enables a more in-depth analysis of their grammatical behavior and collocational patterns, with the aim of creating a usable academic phrase bank for learners.

To answer the research questions, the study employed the tools available in Sketch Engine. First, word sketch difference tool: this tool was used to compare the collocational patterns and grammatical relations of *describe* and *illustrate* across the BAWE and BASE corpora. The query included all inflectional forms of the verbs using the lemma option. A minimum frequency of 1 was set, with the part of speech specified as verb. Collocates were color-coded: green for those exclusive to *describe*, red for *illustrate*, and white for shared collocates. Grammatical relations analyzed included; noun as a subject, noun as an object, and adverb as a modifier. The results were exported to Excel, where collocates were sorted by frequency, log Dice score, and corpus (BAWE or BASE). We then classified noun and adverb collocates by type abstract or concrete noun, adverb of manner or adverb of degree for further analysis.

Second, N-grams tool: to identify common phrase patterns and multi-word expressions, we used the N-gram tool with a length range of 4–6 words. Queries were made in advanced mode using the lemma form of *describe* and *illustrate*, with a minimum frequency threshold of 2. Sample outputs included expressions such as “*can be described as*” or “*clearly illustrate the idea.*” N-gram results were examined for their raw and normalized frequency (per million words) to assess their prevalence within each corpus.

Results and Discussion

RQ1

After we collected all the data both from BAWE and BASE to answer the first research question about the grammatical relations collocational patterns of the verbs *describe* and *illustrate* differ between spoken and written academic English in BAWE and BASE corpora. we analyzed all noun and adverb collocates types that appear as a subject, object, and as a modifier for both words *describe* and *illustrate* in both corpora. We identified each collocates noun and adverb types also by checking its concordance line. In the

complete **table 1** for BAWE there are 100 collocates for describe as a subject and in **table 2** there are 39 noun collocates as a subject in BAWE.

It is important to note that the tables for BAWE showing the results of this word sketch differences are not taken or arranged in the original order of the entire data. Since there are over 100 collocations, we have not included all of them. Therefore, instead, we have selected examples based on color differences to show how certain adverbs are more related to *describe* (in green) and others to *illustrate* (in red) and white to indicate neutral where the colloquial occurs neutrally in both verbs, *describe* and *illustrate*.

Table 1. Word sketch differences result in BAWE for noun modifying a verb as a subject

Collocate	Frequency of Describe	Frequency of Illustrate	logDice Describe	log Dice Illustrate	Noun type
paper	9	0	8,296	--	Concrete
essay	7	0	7,834	--	Concrete
formula	5	0	7,82	--	Abstract
article	11	1	8,699	5,82	Concrete
böll-	5	1	7,915	6,392	Concrete
locke-	4	1	7,336	5,974	Concrete
musil-	3	1	7,183	6,399	Concrete
graph	0	5	--	8,04	Concrete
figure	0	9	--	8,962	Concrete
example	0	14	--	9,673	Abstract

Table 2. Word sketch differences result in BASE for noun modifying a verb as a subject

Collocates	Frequency of Describe	Frequency of Illustrate	logDice Describe	logDice Illustrate	Noun Type
Marx	1	0	7,717	--	Concrete
Right	1	0	7,717	--	Abstract
I	17	4	4,81	2,729	Concrete
Lecture	1	1	7,703	8,164	Abstract
S	1	1	7,952	8,523	
Fact	0	1	--	7,827	Abstract
Increase	0	1	--	10,603	Abstract
Jones	0	1	--	10,999	Concrete

Look at the two tables above, the green color indicates that the word is more associated with *describe* than *illustrate*, and vice versa. The words marked in red are more associated with *illustrate*. Based on the data from the both corpora we obtained, the number of collocates for each also differs due to the significantly different sizes of the corpora. In the BAWE corpus itself, there are 100 collocates, dominated by those associated with *describe*, with 41 strong collocates and 3 middle-strength collocates. As for *illustrate*, there are 6 weak collocates, 36 strong collocates, and 14 neutral collocates that appear with both *describe* and *illustrate*. Therefore, when we compared the overall results from both corpora, the grammatical relation of *a noun modifying a verb as a subject* in the BASE corpus is more dominated by abstract nouns of the 39 collocates found, 26 are abstract nouns and 13 collocates are concrete nouns. In contrast, in the BAWE corpus, concrete nouns are more dominant, with 69 noun collocates out of the total 100 collocates.

However, the five words that appear in both corpora can be considered relatively different because three words such as *study*, *fact*, and *rule* in the BAWE corpus appear with both *describe* and *illustrate*, while *william* and *marx* appear only with *describe*. Conversely, in the BASE corpus, the words *study* and *rule* appear only with *describe*, along with *William* and *Marx*, whereas the word *fact* appears only with *illustrate*.

Table 3. Table of comparing the result of noun modifying a verb as subject from both corpora

BAWE	logDice describe	logDice illustrate	BASE	logDice describe	logDice illustrate
study	6,02	7,32	study	7,9	-
fact	6,0	7,4	fact	-	7,9
rule	5,2	6,9	rule	8,5	-
william	7,5	-	william	8,8	-
marx	7,8	-	marx	7,1	-

Referring to the two tables below, green-highlighted words are more commonly associated with *describe*, while red-highlighted words are more closely linked to *illustrate*. According to the data from the BAWE corpus in **table 3**, there are 100 collocates in total. For *describe*, there are 21 strong collocates and 6 middle-strength collocates. In contrast, *illustrate* has 1 middle-strength collocate and 22 that are considered strong collocates, along with 15 additional weak collocates. Moreover, 35 collocates are neutral, meaning they appear with both *describe* and *illustrate*. When analyzing the grammatical relation where a noun modifying a verb as the object, abstract nouns are predominant, 89 out of 100 while only 9 are concrete nouns. Notably, one noun collocates in BAWE, *structure*, fall into both categories, as their usage in concordance lines it can function as either abstract or concrete nouns depending on the context. Furthermore, it only appears in BAWE but does not appear in BASE.

Table 4. Word sketch differences result in BAWE for noun modifying a verb as an object

Collocate	Frequency of Describe	Frequency of Illustrate	logDice Describe	logDice Illustrate	Noun type
Pain	18	0	8,391	--	abstract
State	11	0	7,388	--	abstract
Method	21	1	7,93	4,062	abstract
Role	12	2	6,985	4,868	abstract
World	13	2	7,695	5,776	concrete
Structure	9	3	6,941	5,996	concrete/abstract
Model	8	5	6,619	6,501	abstract
Example	1	7	3,982	7,563	abstract
Point	2	26	4,773	9,116	abstract
connection	0	4	--	7,241	abstract
possibility	0	5	--	7,314	abstract

Table 5. Concordances examples of concrete and abstract noun for “describe”

Noun type	Left context	KWIC	Right context
Concrete noun	it is useful to	Describe	land <coll>structure</coll> in Nicaragua shortly before the political change.
Abstract noun	Its <coll>structure</col l> is best	Described	as random and fragmented, as Dowell forewarns in chapter 2

In contrast with the result of noun modifying a verb as an object in BASE, with the total 95 noun collocates. Abstract nouns are more dominated than concrete nouns with 12 noun collocates and 83 noun collocates for abstract nouns. 69 noun collocates determined as a strong collocates, 3 noun collocates are neutral, 1 middle-strength collocate, and 22 noun collocates are determined as weak collocates.

Table 6. Word sketch differences result in BASE for noun modifying a verb as an object

Collocate	Frequency of Describe	Frequency of Illustrate	logDice Describe	logDice Illustrate	Noun type
Pain	4	0	9,673	--	abstract
Discomfort	1	0	8,457	--	abstract
Thing	2	1	5,286	4,333	concrete
Sort	3	2	6,025	5,498	abstract
Number	1	1	5,754	5,901	abstract
Point	1	6	5,667	8,384	abstract
Lot	0	1	--	5,227	abstract
Problem	0	1	--	5,359	abstract

However, there are 11 noun collocates that appear in both corpora. But there are significant different in the usage in both corpora. As what can be seen in table below, there are only 2 noun collocates considered as strong collocates in BAWE which are *pain* and *work*. While in BASE, 9 of 11 collocates are considered as strong collocates with *describe* and only 2 noun collocates with *illustrate*. In BAWE, most of the noun collocates are considered as neutral collocates because it appears in both *describe* and *illustrate* such as, *process*, *world*, *experiment*, *function*, *relationship*, *problem*, *difference*. The noun collocate *pain* has the highest logDice score in both corpora.

Table 7. Table of comparing the result of noun modifying a verb as an object from both corpora

BAWE	logDice describe	logDice illustrate	BASE	logDice describe	logDice illustrate
Pain	8,4	-	pain	9,7	-
Work	6,5	-	work	8,2	-
Process	8,4	7,1	process	7,9	-
Type	7,2	4,8	type	6,9	-
World	7,7	5,8	world	7	-
Experiment	7,3	5,6	experiment	7,4	-
Method	8	4	method	7,3	-
Function	6,9	6,1	function	7,3	-
Relationship	8	7	relationship	7,3	-
Problem	7	6,8	problem	-	5,4
Difference	7	7,6	difference	-	6,7

In the several tables above, it can be seen that how *describe* as a verb modify both subject and object. **Table 7** shows the result of word sketch differences for both *describe* and *illustrate* in modifying an adverb as a modifier.

Table 8. Word sketch differences result in BAWE for adverb modifying a verb as a modifier

Collocate	Frequency of Describe	Frequency of Illustrate	logDice Describe	logDice Illustrate	Adverb type
Earlier	15	0	9,446	--	Time
Previously	16	0	8,948	--	Time
Often	14	1	7,432	3,718	Degree
Both	9	2	7,095	5,037	Conjunctive
Here	8	2	8,127	6,414	Place
Fully	7	2	7,637	6,055	Degree
Subsequently	3	1	7,388	6,282	time
Similarly	1	3	5,989	8,135	manner
Clearly	3	31	6,201	9,761	manner
Widely	0	1	--	5,483	manner
Greatly	0	1	--	5,571	degree

It can be seen here, Based on **Table 8**, the adverbs highlighted in green are more frequently used with *describe*, while the adverbs highlighted in red are more frequently used with *illustrate*. From the BAWE corpus, adverbs such as *earlier*, *previously*, and *often* appear more frequently with *describe*, therefore, when viewed they show higher frequencies and logDice scores. On the other hand, adverbs such as *clearly*, *similarly*, and *widely* are more commonly used with *illustrate*. The table also shows the type of adverb for each word, such as time, manner, place, or degree. The adverb of time is most frequently used with *describe*, while the adverb of manner often appears with *illustrate*. From these collocates, it can be seen that *describe* is usually used when talking about the order or time of something, while *illustrate* is used to explain something more clearly or in more detail.

Table 9. Word sketch differences result in BASE for adverb modifying a verb as a modifier

Collocate	Frequency of Describe	Frequency of Illustrate	logDice Describe	logDice Illustrate	Noun type
analytically	1	0	9,475	--	manner
not	3	0	2,202	--	negation
just	9	2	5,762	3,594	degree
here	5	3	6,252	5,542	place
also	3	2	5,71	5,146	conjunctive
actually	0	1	--	3,352	conjunctive
well	0	1	--	3,74	manner
slightly	0	1	--	7,102	degree

Table 9 shows *describe* and *illustrate* as verbs modify as a modifier from the BASE corpus. Adverbs such as *analytically*, *not*, and *just* occur more frequently with *describe*, while *actually*, *well*, and *slightly* occur more frequently with *illustrate*. Therefore, it can be interpreted that in spoken academic English, *describe* is usually used with adverbs that indicate manner, negation, and place. On the other hand, *illustrate* often occurs with adverbs that indicate degree or strengthen a meaning. Compared to the BAWE corpus, which is based on written text, BASE shows how these verbs are used in spoken language.

Table 10. Table of comparing the result of for adverb modifying a verb as a modifier from both corpora

BAWE	logDice describe	logDice illustrate	BASE	logDice describe	logDice illustrate
Already	6,7	-	already	5,5	-
sometimes	7,9	-	sometimes	7,6	-
Indeed	5,6	6	indeed	7,6	-
Exactly	6,9	6,3	exactly	7,6	-
Just	6,7	-	just	5,8	3,6
Here	8,1	6,4	here	6,2	5,4
Also	6,4	5,9	also	5,7	5,1
Well	5,3	7,7	well	-	3,7

In the comparison table between BAWE and BASE above, there are collocates that appears in both corpora. However, the logDice scores are different, and the difference also appear in *indeed*, where in BASE it considers as a neutral while in BAWE is considered as green or related more to the verb describe. This table also shows that both *describe* and *illustrate* have some kind of a similar pattern in both spoken and written academic English. The difference is just in the strength for each collocates because it is depending on the corpus.

From the findings and discussion for the first research question above, In BAWE corpus, *describe* frequently co-occurs with concrete noun subjects such as *paper*, *essay*, and etc. While *illustrate* is linked to visual or abstract elements such as *figure* and *example*. These results aligned with Liu (2023) that did the comparative study on English synonym for illustrate, explain, interpret and found that *illustrate* often collocated with nouns representing *example*, *charts*, *diagram*, and *graph*. In BASE, abstract nouns dominate as subjects for both verbs. As an object, abstract nouns are more common in both corpora, but BAWE features a large range, including context-dependent nouns where it can be concrete or abstract nounlike *structure*. Adverb collocates also differ in both corpora. *Describe* in BAWE is often used with time-related adverbs such as *earlier* and *previously*, while *illustrate* is paired with adverbs of manner and emphasis like *clearly* and *similarly*. This is once again also aligned with Liu (2013) where in his study using COCA Corpus *clearly* is the highest adverbial collocations with the frequency 296 and the MI score 3.57. In BASE, describe tends to co-occur with manner and place adverbs for the example like *just*, and *here*. While illustrate appears with degree modifiers like *well* and *slightly*. It can be confirmed that RQ1 has been answered.

RQ2

The second research question focusing on what kinds of phrases that are typically collocates with *describe* and *illustrate* in academic spoken and written texts. In order to answer this research question, we used the data we got from the n-gram sketch engine tool. We matched the n-gram data with the collocations found from the previous research question and organized them into a phrase pattern table. This table highlights common academic phrases in both spoken and written English, and it also provides information such as the noun and adverb types of the collocations. This allows students to see how describe and illustrate are used in spoken and written context. It is crucial to note that the phrases in this table are not whole phrase generated by the N-grams tool. Based on RQ1 findings, we chose the parts that collocate directly with nouns or adverbs from the word sketch differences tool result both in BAWE and BASE Corpora.

Table 11. Table of N-Grams results in BAWE for *describe*

N-grams (4-6) result for <i>describe</i>	F	Relative frequency/million	collocates	POS of the collocate
describe above, the...	11	1,32	above	Adverb
be often describe as...	11	1,32	often	Adverb
also be describe as...	11	1,32	also	Adverb
be describe below...	9	1,08	below	Adverb
describe the relationship between...	6	0,72	relationship	Noun
describe the behaviour of...	6	0,72	behaviour	Noun
describe the use of...	6	0,72	use	Noun
,as describe earlier...	5	0,60	earlier	Adverb
this paper describe the...	5	0,60	paper	Noun
describe the effect of...	4	0,48	effect	Noun
describe the fact that...	4	0,48	fact	Noun
describe the process of...	4	0,48	process	Noun
be often describe as a...	3	0,36	often	Adverb
can also be describe as a...	3	0,36	also	Adverb
often describe as...	3	0,36	often	Adverb
can therefore be describe...	3	0,36	therefore	Adverb
as describe earlier...	3	0,36	earlier	Adverb

Table 11 above shows the result of N-grams tool in sketch engine with the N-grams length 4-6 that include the verb *describe* in academic texts in BAWE Corpus, along with their frequency, relative frequency per million words, common collocates, and its part of speech (noun or adverb). The data reveals that *describe* often appears with adverbs such as *above*, *often*, *also*, *below*, and *earlier*. For example, phrases like *be often describe as...* or *as describe below...* show that *describe* is commonly used in passive forms and helps writers refer to information already mentioned or explained later in the text.

In addition, *describe* frequently co-occurs with nouns such as *relationship*, *behaviour*, *use*, and *process*. which is categorise as an abstract noun. Expressions like *describe the relationship between...* or *describe the process of...* suggest that this verb is used to explain key concepts, actions, or facts. These patterns indicate that *describe* serves two main purposes in academic writing. The first one is to point to or highlight certain parts of the text often with the help of adverbs, and the second is to explain important topics or ideas whether with nouns as objects.

Table 12. Table of N-Grams results in BAWE for *illustrate*

N-grams (4-6) result for <i>illustrate</i>	F	Relative frequency/million	collocates	POS of the collocate
illustrate the importance of...	17	2,04	importance	Noun
be illustrate below...	10	1,20	below	Adverb
be best illustrate by...	5	0,60	best	Adverb
to illustrate the importance...	5	0,60	importance	Noun
to illustrate the importance of...	5	0,60	importance	Noun
to illustrate this point...	5	0,60	point	Noun
be well illustrate by...	5	0,60	well	Adverb
to illustrate this point...	4	0,48	point	Noun
,as illustrate below...	4	0,48	below	Adverb
to illustrate your argument...	4	0,48	argument	Noun

illustrate your argument...	3	0,36	argument	Noun
be illustrate below, figure...	3	0,36	below	Adverb
,this be well illustrate...	3	0,36	well	Adverb
this be best illustrate by...	3	0,36	best	Adverb
this be best illustrate...	3	0,36	best	Adverb
,this be well illustrate by...	3	0,36	well	Adverb
this be well illustrate by...	3	0,36	well	Adverb
this be well illustrate...	3	0,36	well	Adverb
,clearly illustrate the...	3	0,36	clearly	Adverb
,to illustrate further...	3	0,36	further	Adverb
to illustrate your argument...	3	0,36	argument	Noun
clearly illustrate that the...	3	0,36	clearly	Adverb
illustrate the ha result...	3	0,36	result	Noun
illustrate the ha result for...	3	0,36	result	Noun

Table 12 above shows that *illustrate* appears with abstract nouns such as *importance*, *point*, *argument*, and *result*. These nouns typically function as objects of the verb. *Illustrate* is commonly used to explain research findings or reasoning. The use of *importance* and *argument* repeatedly indicates that *illustrate* plays a role in supporting claims or highlighting the significance of certain aspects in a discussion session especially in academic writing. In addition, *illustrate* is often modified by adverbs such as *below*, *best*, *well*, *clearly*, and *further*. These adverbs help specify how, how well, or where the illustration occurs.

Table 13. Table of N-Grams results in BASE for *describe*

N-grams (4-6) result for <i>describe</i>	F	Relative frequency/million
be describe in term...	3	1,97
and that be describe...	3	1,97
be describe in term of...	3	1,97
i have just describe...	3	1,97
can be describe by...	3	1,97
that i have describe...	3	1,97
try to describe the...	3	1,97

The N-grams data presented in the BASE corpus for the verb *describe* in Table 14 shows a clear difference from the results found in the BAWE corpus. The phrases listed in BASE such as *be describe in term of...*, *can be describe by...*, and *try to describe the...* are not only fewer in number but also less varied in structure and word choice when it compares with the result from BAWE Corpus. This table actually shows the entire set of available N-gram results for *describe* in BASE, while BAWE produced more than 300 different N-gram phrases or multiword expressions (MWEs) for the same verb. This strong contrast highlights both the lower frequency and the limited variety of *describe* in spoken academic contexts.

In BAWE, *describe* often appears with a wide range of noun and adverb collocates for the example *relationship*, *process*, *above*, and *clearly* that many of which were also found in the RQ1 results.

Table 14. Table of N-Grams results in BASE for *illustrate*

N-grams (4-6) result for <i>illustrate</i>	F	Relative frequency/million
be illustrate in the...	3	1,97

As it can be seen in **Table 14**, The N-gram result for *illustrate* in the BASE corpus shows only one phrase *be illustrate in the...*, with a frequency of 3. This shows that *illustrate* is not often used in long or specific phrases in spoken academic English. Compared to the BAWE corpus in **Table 13**, where *illustrate* appears in many academic phrases like *illustrate the importance of...* or *clearly illustrate the...*, the BASE result is very limited.

This limited result from BASE may be caused by a few factors. First, spoken academic language is often not as formal as written language. Second, *describe* itself might not be used as frequently in spoken academic settings compared to written ones, where writers need to give more detailed explanations. Another important reason that also play a significant role on the results is the difference in corpus size where BAWE is larger than BASE. Because of this, we cannot clearly see how the collocates found in RQ1 are linked to the phrases shown in the BASE N-gram results. In BAWE, many collocates as the result of word sketch difference tool before exist in one phrase with *describe* (*importance*, *argument*, *fact*, *result*, *process*, and etc.), like *describe the importance of...* or *be often describe as*. But in BASE, the results are too few and simple to make a direct connection. Therefore, it can be concluded that RQ2 has been answered, showing that the types of phrases typically collocated with *describe* and *illustrate* strongly differ between written and spoken academic contexts.

Conclusion

This study examines descriptive verbs, namely *describe* and *illustrate* from two academic corpora, namely BAWE and BASE. The results of the first research question showed that *describe* tends to collocate with concrete nouns which are *paper*, *essay*, and *article* while *illustrate* tends to collocate with abstract nouns such as *example* and *figure*. However, there is also something called context-dependent where the collocates can enter into concrete or abstract nouns depending on the context of the existing concordances lines. such as *structure*. The patterns obtained from the BASE corpus data are simpler and less varied, with fewer academic collocations found for both verbs.

For the second research question, the N-grams tool analysis shows very large results for BAWE but very little for BASE. However, the BASE result of N-grams tools does not include these academic collocates and even it shows simpler and more common phrases. For example, phrases like *I have just describe...* or *that I have describe...* reflect a more personal and spoken style, where the speaker focuses more on sharing their thoughts than on giving detailed or formal explanations. Therefore, due to the smaller corpus size and informal tone of academic speech. For the future research are encouraged to do the research beyond more than only two descriptive verbs for enriching the data. Researcher should consider the corpus size and consider to use larger corpus in order to get more data. especially in spoken academic English. Researcher could also do the research representing non-native English Speakers and identifying their collocational patterns due to the difference background and culture. It might be valuable insights especially in ESL or EFL ultimately in academic contexts.

References

- Chaengchenkit, R. (2023). A corpus-based study of the synonyms ceases, halt, and stop. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 473–494. <https://so04.tci-thaijo.org/index.php/LEARN/index>
- Chon, Y. V., & Shin, D. (2013). A corpus-driven analysis of spoken and written academic collocations. *Multimedia-Assisted Language Learning*, 16(3), 11–38. <https://doi.org/10.15702/MALL.2013.16.3.11>
- Hong, A. L. (2019). A CORPUS ANALYSIS OF PHRASEOLOGICAL SEQUENCES IN ACADEMIC WRITING. *The European Proceedings of Social & Behavioural Sciences*, 413–418. <https://doi.org/10.15405/epsbs.2019.09.46>
- Jarrar, Y. A., Al-Sayyed, S. W., Rababa'h, B. B., et al. (2024). A corpus-based analysis of four near-synonymous English verbs. *Forum for Linguistic Studies*, 6(6), 678–701. <https://doi.org/10.30564/fls.v6i6.7378>
- Liu, S. (2023). A corpus-based comparative study on English synonyms in COCA: Taking illustrate, explain, and interpret as examples. *Hill Publishing*, 7(5), 577–580. <https://doi.org/10.26855/er.2023.05.007>
- Narkprom, N. (2024). A corpus-based investigation into English synonymous verbs restrict and constrain. *LEARN Journal: Language Education and Acquisition Research Network*, 17(1), 688–714. <https://so04.tci-thaijo.org/index.php/LEARN/index>
- Nesi, H., & Gardner, S. (2018). The BAWE corpus and genre family's classification of assessed student writing. *Assessing Writing*, 1–5. <https://doi.org/10.1016/j.asw.2018.06.005>
- Orhon, Y., Kulac-Puren, D., & Guzel, E. (2018). A corpus linguistic study on the use of however in British academic spoken and written English. *The Literacy Trek*, 4(2), 31–62. <https://dergipark.org.tr/en/pub/literacytrek/issue/41278/401984>
- Pratiwi, D. D. (2020). Frequency and collocation of the verbs refuse and reject. *UC Journal: ELT, Linguistics and Literature Journal*, 1(2), 150–182. <https://doi.org/10.24071/uc.v1i2.3020>
- Selmistraitis, L. (2020). Semantic preference, prosody and distribution of synonymous adjectives in COCA. *GEMA Online Journal of Language Studies*, 20(3), 1–18. <https://doi.org/10.17576/gema-2020-2003-01>
- Szczygłowska, T. (2019). A corpus-based study of the specificity adjectives specific and particular in academic written English: evidence from the BAWE corpus. *Brno studies in English*, 45(2), 175–199. <https://doi.org/10.5817/BSE2019-2-9>