European Journal of English Language Teaching
ISSN: 2501-7136
ISSN-L: 2501-7136
Available on-line at: www.oapub.org/edu

DOI: 10.46827/ejel.v9i1.5202 Volume 9 | Issue 1 | 2024

LEXICAL COLLOCATIONS OF KEYWORDS USED IN BUSINESS NEWS: A CORPUS-BASED STUDY

Krittat Sukman[i],
Supong Tangkiengsirisin
Thammasat University,
Thailand

Abstract:
This corpus-based study aims to facilitate the instruction of English for Specific Purposes (ESP) and benefit Thai undergraduate students majoring in Business English by investigating lexical collocations found in business news. A corpus of business news (COBN) was compiled, containing 678,373 running words sourced from 755 online business news. Employing AntConc version 3.5.9, the first 120 keywords were identified in accordance with the set criteria, with the majority being nouns (82.5%), followed by adjectives (12.5%), and verbs (5%). These keywords served as "nodes" to identify their "collocates," resulting in 330 pairs of lexical collocations with 5 combination types. Following a framework adapted from Benson, Benson, and Ilson, prominent combinations included Noun + Noun (54.02%), Adjective + Noun (26.86%), Noun + Verb (13.69%), Verb + Noun (4.47%), and Verb + Adverb (0.89%). The pedagogical implications and recommendations for future research were also discussed.

Keywords: business news, collocations, corpus analysis, keyword lists

1. Introduction

Collocations refer to sets of two or more words that occur together more frequently than expected based on individual frequencies (Jones & Sinclair, 1974). They are considered a crucial component of lexical development in a second language, contributing significantly to communicative competence (Henriksen, 2013). Mastery of prefabricated chunks or formulaic sequences, with collocation as a key aspect, is emphasized for effective language use (Henriksen, 2013). Hill (2000) and Nation (2001) similarly underscore the essential role of collocation in achieving native-like fluency. That is, native speakers and proficient language users inherently grasp word co-occurrences. Therefore, language teachers are urged to incorporate collocation instruction to enable learners to become competent users of the target language.

[i] Correspondence: email krittatt_s@hotmail.com

In the realm of English for Specific Purposes (ESP), the significance of understanding collocations cannot be emphasised enough. It is crucial for ESP learners to grasp domain-specific lexicon, particularly combinations that can alter the original meaning, particularly in their specific fields like business, accounting, economics, nursing, tourism and hospitality, medical sciences, engineering, etc. Önder Özdemi, 2014; Tabak and Takač, 2023). This holds true for Thai business English students, especially those under examination in this study. Throughout their four-year undergraduate programme in Business English, students are required to engage with numerous business news articles. To enhance their comprehension of current business news, it is essential for them to acquire awareness of the collocations used in these articles. Despite the inclusion of business collocations in classroom coursebooks, they lack corpus-based information, relying instead on authors' intuition. Consequently, learning collocations specific to their professional discourse, such as those found in business news, is imperative for both academic success and future professional communication.

Over the past decades, language teaching professionals have observed the growing processing power of computer technology, enabling the convenient compilation of frequency lists for lexical items using corpus software. Corpus-based approaches have been employed in various studies to identify high-frequency collocations or specialised vocabulary in different disciplines (e.g., Menon and Mukundan, 2012; Ackermann and Chen, 2013; Parkinson, 2015; Kiss and Horváth, 2015; Chen, 2017; Trinant & Yodkamlue, 2019; Phoocharoensil, 2020; Trinant and Kijpoonphol, 2021; Otto, 2021; Suraprajit, 2022; Kadirbekova, 2023). However, there is a notable gap in the study of collocations specific to online business news, presenting an unexplored avenue. The paucity of literature in this area means that collocation lists for online business news remain elusive.

Consequently, this present study aims to investigate the keywords essential for business English students when reading business news articles and identify the most frequent collocations found in these online business news articles.

To achieve these objectives, the study addresses the following research questions:

1) What are the keywords found in the corpus of online business news?
2) What are the lexical collocations of the keywords from the corpus of online business news?

2. Literature review

2.1 Vocabulary types and keywords

Nation (2001) categorises words into four groups, namely high-frequency words, academic vocabulary, technical vocabulary, and low-frequency words. High-frequency words, constituting the top 2,000 most commonly used words in English or the General Service List (GSL), encompass around 80% of running words in any text (West, 1953). Academic words, derived from the Academic Word List (AWL) by Coxhead (2000), consist of 570 word families commonly found in academic texts, constituting approximately 10% of running words. Technical words, absent in GSL and AWL, are

specific to certain subject areas, representing about 5% of the text. Lastly, low-frequency words cover only a minimal proportion of the texts.

Nation's classification is valuable because 'key' words should not fall into the high-frequency group. A keyword is defined as "*a word which occurs with unusual frequency in a given text*" (Scott, 2012, p. 237) compared to its occurrences in a reference corpus. Essentially, a keyword has a higher frequency than expected by chance compared to a reference corpus. In corpus linguistics, the extraction of keywords is a commonly used tool alongside frequency profiling, concordance, n-grams (clusters or lexical bundles), and collocation analysis. Various studies in linguistics and other disciplines have utilised the keyword procedure for textual data analysis, ranging from a discourse analysis of refugees and asylum seekers in the UK press, studies of health communication, and lexical simplification in translations, to profiling learners' language and developing e-learning materials.

2.2 Collocations

As previously noted, the most effective way to acquire L2 vocabulary is through the learning of frequently occurring and typical collocational patterns. Jones and Sinclair (1974) define collocation as sets of two or more words that appear together more often than would be expected based on their individual frequencies. For example, "night" is expected to frequently collocate with "dark." Further, Schmitt (2000) provides another corresponding definition of collocation, describing it as "*the tendency of two or more words to co-occur in discourse*" (p. 76). He emphasises understanding collocations through their co-occurring nature and varying degrees of exclusivity. For example, the word 'blonde' may exclusively collocate with 'hair' rather than with words like 'paint' or 'wallpaper.' Bennett (2010) also defines collocation as "*the statistical tendency of words to co-occur*" (p. 8). She explains that if one word is used, there is a high statistical probability of certain words co-occurring. At this point, it is reasonable to conclude that collocation involves the highly predictable co-occurrence of words, as native speakers tend to select them together with statistical significance in terms of their association. For this reason, McCarthy and O'Dell (2005) endorse the idea that learning collocations can enhance the ability to communicate in English more naturally and accurately, both in spoken and written forms.

Among various classifications, Benson *et al*. (2010) emerge as the most well-known and particularly relevant in this study. They categorise collocations into two main types: lexical collocations and grammatical collocations. Lewis (2000) and Newman (1988) consider lexical collocations as the most common type in the English language, making it a focal point in this study. This type typically comprises two or more content words, such as nouns, verbs, adjectives, and adverbs.

In this study, the researcher adheres to the classifications presented by Benson e*t al*. (2010) due to their clear elucidation of collocation types, aligning with the study's objectives.

Table 1: Lexical collocations as categorised by Benson *et al.* (2010)

Types	Combinations	Examples
1	Verb + Noun	do a business, launch a product
2	Adjective + Noun	reckless abandon, rough estimate
3	Noun + Verb	bombs explode, blood circulates
4	Noun + Noun	wound dressing, road accident
5	Adverb + Adjective	heavily influenced, bitterly disappointed
6	Verb + Adverb	hurt badly, increase dramatically

2.3 Relevant studies

Several previous studies have been carried out to explore lexical collocations in various text types. For instance, Ackermann and Chen (2013) created the Academic Collocation List (ACL) by constructing a corpus that included 28 academic disciplines sourced from journal articles and textbook chapters. The corpus, totaling 25.6 million words, yielded 2,468 entries of lexical collocations. The predominant combinations were Adjective + Noun, Verb + Noun, and Adverb + Adjective. In Önder Özdemir's (2014) study, a small-specialised corpus revealed useful collocations for learning Medical English such as *administer treatment* and *tentative diagnosis*. Also, Kiss and Horváth (2015) conducted a study aimed at enhancing the teaching and learning of tourism vocabulary using the content-integrated language learning (CLIL) method in a Hungarian vocational school. The project involved creating a corpus of beach holiday-related articles, totaling 51,150 words.

The vocabulary analysis, utilising the AntConc Keyword function's Chi-square test, identified eight content keywords among the top 100 in the corpus, including beach, island, town, bay, ocean, coast, resort, and sea. The subsequent examination of word spans on the left revealed collocates of the first two keywords, highlighting evaluative and descriptive adjective-noun clusters based on frequency and statistical measures. Otto (2021) created two corpora of civil engineering to improve the teaching and learning of specialised vocabulary for civil engineering students. To identify important specialised vocabulary, keywords analysis was first performed by using AntConc with log-likelihood ratio with COCA being used as a reference corpus. The findings revealed 18 specialized words and their collocates. More recently, Kadirbekova (2023) constructed a specialised Information-Communication Technologies (ICT) corpus to aid instructors in identifying collocations for communication, report writing, and interpretation in the field. With the application of AntConc software, the study employed the Word List function to identify frequently used collocations in the ESP corpus related to ICT. The analysis revealed the top 10 collocations in ICT journals, such as *software engineering, cognitive radio, computer science, cyber (and) physical attack, cyber risk, cyberattack, software systems, cyber security*, *software development*, and *radio networks*. The researcher highlighted the importance of specialised corpora for understanding professional language and assisting ESP teachers in comparing results with textbooks and other materials.

In Thailand, Trinant and Yodkamlue (2019) conducted a study to identify lexical collocations in nursing to enhance the teaching and learning of nursing English. They compiled a 1.25-million-word corpus from 300 research articles in 10 nursing journals.

Using AntConc version 3.4.4, they identified 717 keywords, primarily nouns (63.51%), adjectives (21.54%), verbs (13.44%), and adverbs (1.51%). These keywords served as "nodes" to find their "collocates," resulting in 2,148 pairs of lexical collocations. The majority were Noun-Noun (41.39%), Adjective-Noun (28.4%), and Noun-Verb (11.17%). They applied a Mutual Information (MI) score of 5 or higher to measure word association strength, considering the frequency of co-occurrence. Similarly, Trinant and Kijpoonphol (2021) endeavoured to investigate lexical collocations of keywords by constructing a tourism research article corpus, encompassing 1.5 million running words from 240 articles in six tourism and hospitality management journals. The keywords were predominantly nouns (62.72%), followed by adjectives (18.37%), verbs (16.37%), and adverbs (2.53%). These keywords served as "nodes" to identify their "collocates," resulting in 2,989 pairs of lexical collocations with 16 combination types. Notably, Noun + Noun (44.76%), Adjective + Noun (25.89%), and Noun + Verb (9.33%) were the most prevalent combination types, aligning with a framework adapted from Benson *et al*. (2010).

Additionally, Suraprajit (2022) conducted a study analyzing the English language used in online logistics magazines in Thailand. Using the corpus tool Ant-Conc, the study examined the frequency of nouns, verbs, adjectives, and collocations in these texts. The results highlighted the ten most common nouns, verbs, and adjectives, as well as the dominant types of grammatical and lexical collocations. Notably, the findings have potential implications for English for Specific Purposes (ESP) research and underscore the value of incorporating authentic language from the logistics business in language instruction.

Since this study was aimed at facilitating the learning and teaching of lexical collocations specifically for Thai business English students, the Corpus of Online Business News (COBN) was created. The study focused on identifying and classifying lexical collocations within the corpus, using keywords as 'nodes' to find their 'collocates.' Given that lexical collocations involve the co-occurrence of two content words close to single-word vocabulary, the primary emphasis was on this type of collocation.

3. Method

3.1 Research instruments

3.1.1 Online business news

A total of 755 business news articles from BBC.com, spanning the years 2021 to 2022, were deliberately chosen for analysis. This selection was motivated by the intention to expose students to the most recent keywords and collocations essential for comprehending business news. The complete news articles were preserved in their original length and electronically stored in a corpus referred to as COBN (Corpus of Online Business News). However, certain non-article content, such as charts, diagrams, tables, numbers, references, and photographs, was excluded. Regarding the business news section on the web, several domains of business are addressed in different topics such as business, market data, new economy, new tech economy, companies, entrepreneurship, and technology of business. The rationale behind this is that these business news articles are

important for students majoring in Business English since they serve as both their required reading materials throughout the course of their study.

3.1.2 Corpus analysis tool

In this study *AntConc 3.5.9* (Anthony, 2020) was employed as a tool for analysing the compiled corpus, revealing 678,373 running word tokens and 23,436 word types within the scope of AntConc analysis. It is important to acknowledge that COBN, while a small corpus, serves the purpose of representing the English language used in business news articles. The choice of AntConc as a corpus analysis tool is justified for several reasons. Firstly, it is a freely available and versatile tool suitable for corpus linguistics research and data-driven learning. Its compatibility with various operating systems, including Microsoft Windows, Macintosh OS X, and Linux, enhances its accessibility. Additionally, AntConc offers a range of seven corpus tools, such as Concordance Tool, Concordance Plot Tool, File View Tool, Cluster/N-grams, Collocates, Wordlist, and Keyword List, aligning with the objectives of the current study. Furthermore, previous research, including studies by Hou (2014), Kiss and Horváth (2015), and Trinant and Yodkamlue (2019), has demonstrated the successful application of AntConc in similar analyses.

3.2 Research procedure

3.2.1 Identifying keywords

To identify lexical collocations in the business news corpus, a preliminary step involves determining keywords to serve as nodes for identifying their collocates. The British National Corpus (BNC) was employed as a reference corpus due to its extensive use in various studies. The creation of the keyword list involves deciding the number of keywords based on keyness value and frequency of occurrence. Following Baker's (2006) guidance, a keyness score of ≥ 100 was set for this study. Additionally, a frequency value of ≥ 50 was established. It is crucial to note that certain word types, such as pronouns, proper nouns, function words, abbreviations, and acronyms, were excluded from the keyword list. From this process, only the top 120 keywords were selected for collocate analysis, considering the infeasibility of analysing all generated keywords. Notably, only keywords with a specialised meaning related to business disciplines, determined through a "*rating scale approach*" (Chung and Nation, 2004), were chosen for analysis. The rating process involved manual evaluation by the researcher and experts in the field of business.

3.2.2 Identifying lexical collocations

After the keywords were determined, they were used as the 'nodes' to further identify their collocates. The criteria for collocation identification were as follows:

1) The range of the collocates were set within the 3-word span on the right of the node. This was intended to explore the collocates that occurred after the nodes within this range of word span due the reason that this range was neither too close nor too far for each pair to co-occur.
2) Only the lexical collocations were chosen.

3) Then the strength of each pair of collocations was measured by Mutual Information (MI) as provided by AntConc. In this study, the MI score was set at ≥ 5.
4) The frequency of the cooccurrence of each pair was set at ≥ 10.
5) In case there were no applicable collocates with the MI score of ≥ 5, the collocates with the highest MI value would be selected.
6) If no suitable collocates were identified, even with an MI value of < 5, a reduction in the frequency of occurrence would be implemented.

4. Results

4.1 The list of top 120 keywords

With the determined criteria where with frequency of occurrence is at ≥ 50 and the keyness value is ≥ 100, a total of 2,942 keywords were generated. However, there are some words other than content words appearing on the keyword list that are not the focus of this present study and thus deleted; namely, function words, pronouns, proper nouns, reporting verbs, and abbreviations. Also, for the sake of manageability and time constraints, only the first 120 keywords were opted to study their collocations. Therefore, the keywords found in COBN that are used as "nodes" were 120. Of this number, 'pandemic' has the highest keyness value, which is 8807.01, whilst the lowest keyness value belongs to 'production' at 264.51.

The list of the first 120 keywords is presented in Table 1 below.

Table 1: The top 120 keywords found in COBN

No	Freq.	Keyness	Keywords	No.	Freq.	Keyness	Keywords
1	916	8807.01	pandemic	61	211	605.82	investors
2	1046	3590.49	prices	62	209	588.66	giant
3	467	3131.75	online	63	227	581.93	delivery
4	993	2872.3	energy	64	76	579.89	startup
5	632	2714.66	global	65	535	577.93	hit
6	1428	2267.3	company	66	187	572.49	petrol
7	942	2090.86	companies	67	230	550.47	chain
8	451	1689.03	businesses	68	155	534.45	emissions
9	699	1650.87	firm	69	166	530.57	households
10	648	1590.81	oil	70	198	530.19	bills
11	472	1583.93	inflation	71	120	530.14	shortage
12	157	1516.95	lockdown	72	313	524.91	media
13	544	1510.93	gas	73	215	519.11	electric
14	1076	1459.28	business	74	99	518.78	streaming
15	184	1438.05	internet	75	384	498.53	growth
16	133	1303.22	lockdowns	76	151	486.52	restaurants
17	416	1267.2	rising	77	434	483.52	financial
18	129	1264.03	cryptocurrency	78	178	456.8	restrictions
19	431	1209.09	employees	79	606	449.08	market
20	543	1198.78	economy	80	242	434.63	crisis
21	122	1195.43	website	81	94	424.83	soaring

22	443	1185.78	customers	82	254	417.83	users
23	259	1170.4	sanctions	83	99	414.38	hybrid
24	378	1168.65	fuel	84	123	408.63	flights
25	697	1161.83	food	85	127	392.8	shortage
26	698	1156.6	price	86	140	387.7	imports
27	173	1140.37	spokesperson	87	122	386.43	crude
28	456	1123.02	firms	88	278	368.11	jobs
29	608	1103.77	workers	89	166	360.58	employee
30	279	1101.78	invasion	90	192	357.04	consumer
31	272	1043.63	digital	91	105	356.83	infrastructure
32	689	1030.11	pay	92	190	356.63	store
33	487	1001.11	sales	93	94	356.73	wholesale
34	203	1000.43	platforms	94	377	349.6	higher
35	194	983.91	retailers	95	70	347.35	regulators
36	450	908.91	supply	96	368	346.55	living
37	264	894.39	platform	97	263	328.17	sector
38	391	862.25	impact	98	368	325.67	increase
39	240	860.66	retail	99	86	324.67	retailer
40	450	818.36	demand	100	155	324.13	supplies
41	206	806.43	founder	101	87	323.15	sustainable
42	391	794.3	executive	102	88	307.36	analyst
43	167	778.37	brands	103	143	300.98	remote
44	429	769.19	products	104	453	295.22	office
45	211	767.26	rises	105	54	293.79	windfall
46	588	755.28	industry	106	54	292.62	sustainability
47	439	738.52	chief	107	88	286.72	owns
48	75	734.89	blockchain	108	94	280.62	analysts
49	728	734.79	working	109	54	280.54	logistics
50	497	722.96	costs	110	333	279.34	started
51	221	720.44	stores	111	44	279.17	disruptions
52	401	673.2	rise	112	69	277.62	soared
53	211	664.5	consumers	113	32	276.95	startups
54	585	663.02	staff	114	101	276.87	shipping
55	506	648.61	bank	115	83	275.35	workplace
56	419	645.59	technology	116	58	274.06	regulator
57	569	643.31	cost	117	72	271.17	bookings
58	523	638.97	data	118	68	268.12	deliveries
59	193	610.43	brand	119	187	266.14	profit
60	206	608.29	drivers	120	329	264.56	production

Among these keywords, the vast majority of them are nouns (N = 99), accounted for 82.5 percent. The adjectives (N = 15) came second and accounted for 12.5 percent. The smallest number among them are verbs (N = 6), merely 5 percent.

A detailed illustration of each part of speech of the keywords and their percentage is provided below:

Table 2: Keywords generated from COBN according to their parts of speech

No.	Parts of Speech	Numbers	Percentage
1	Noun	99	82.5
2	Adjective	15	12.5
3	Verb	6	5
	Total	**120**	**100**

The dominant occurrence of content words like nouns, adjectives, and verbs in the keywords appears to be a shared characteristic across all corpora. This observation aligns with similar findings in studies conducted by other researchers, including Coxhead's Academic Word List (2000), Ackermann and Chen (2013), Kiss and Horváth (2015), Trinant and Yodkamlue (2019), Trinant and Kijpoonphol (2021), Suraprajit (2022), and Kadirbekova (2023).

4.2 Lexical collocations in COBN

After the keywords from the COBN had been identified, the process of collocation extraction commenced. which followed the set criteria, namely the word-span of 3 on the right side of the node (3R), the set frequency of occurrence, and the mutual information (MI) score.

The summary of the findings is provided in Table 3 below.

Table 3: Lexical Collocations in COBN according to combination types

No.	Combination Types	No. of Collocation Pairs	(%)
1	Noun + Noun	181	54.02
2	Adjective + Noun	90	26.86
3	Noun + Verb	46	13.69
4	Verb + Noun	15	4.47
5	Verb + Adverb	3	0.89
	Total	**330**	**100**

Table 3 above presents 330 collocation pairs extracted from the 120 keywords identified earlier from the corpus. It is apparent that most combination pairs are the 'Noun + Noun' combination with 181 pairs that account for over 54.02 percent of all combination types. The 'Adjective + Noun' combination ranks second with 90 pairs (26.86%), followed by the 'Noun + Verb' combination which makes 46 pairs (13.69%). The 'Verb + Adverb' combination generates 15 pairs (4.47%), and the 'Verb + Noun' combination has the least number with three pairs being generated (0.89%).

Examples of collocation pairs of each combination type are illustrated in Table 4 below.

Table 4: Examples of Lexical Collocations extracted from COBN

Nodes	Collocates	Examples
Noun	Noun	pandemic restrictions, petrol (and) deisel, growth forcast, production facilities, supply shortages, rise (in) cost, price cap
	Verb	pandemic hit, prices surged, economy rebounded, restrictions (were) lifted, sanction imposed, bills (were) set
Adjective	Noun	online sales, chain disruption, wholesale prices, hybrid workers, financial stability, global economy, digital assets
Verb	Noun	pay attention, hit record, increase productivity
	Adverb	(is)working remotely, hit hard

When considering the nodes or keywords with their collocates, it is interesting to discover the nodes that attract the most collocates. The keywords/nodes of each part of speech above thus have been identified. Particularly, the five noun keywords with the most collocates include: *prices*, *energy*, *gas*, s*upply*, and *food*. The five adjective keywords the most collocates are: *global*, *financial*, *rising*, *online*, and *chief*. Also, of all the four verb keywords, *increase* appears to be the node with the most collocates.

The details are presented in the following in Tables 5 to 7.

Table 5: Five noun keywords with the most collocates

Keywords (No.)	Nouns	Verbs
energy (16)	bills, cap, bill, consumption, transition, costs, strategy, agency, prices, suppliers, security, crisis, supplies, giant, price, markets	-
gas (9)	pipeline, electricity, imports, exports, producers, supplies, emissions, bills, prices	-
price (9)	record, index	surged, risen, jumped, soared rise, rising, rose
supply (8)	chains, chain, disruptions, issues, shortages, problems, crisis, demand	-
food (8)	waste, drink, delivery, items, prices, fuel, inflation, costs	-

Table 6: Five adjective keywords with the most collocates

Keywords (No.)	Nouns	Adjectives
global (15)	benchmark, shortage, chip, chains, supply, economy, markets, crisis, economic, trade, growth, market, production, energy, oil	economic, financial
financial (10)	crime, institutions, stability, officer, system, crisis, times, markets, services, support	-
rising (7)	costs, prices, inflation, fuel, bills, energy, food	-
online (6)	shopping, retailer, orders, sales, platforms, platform	-
wholesale (6)	prices, petrol, gas, costs, price, fuel	-

Table 7: The verb keyword with the most collocates

Keywords (No.)	Nouns
increase (4)	productivity, number, production, demand

It is noticeable examining the above Tables 5 to 7 that noun nodes/keywords are more likely to co-occur with noun collocates and some tendency to attract verb collocates. The adjective nodes are more common to collocate with nouns, albeit some chance to co-occur with adjective collocates. The node verbs, however, have a great possibility of co-occurrence with noun collocates and may collocate with adverbs. When ranking the lexical collocations according to their frequency of occurrence, it is clear that the 50 most frequent collocations are Noun + Noun and Adjective + Noun combinations. Accordingly, it should be beneficial that the test of collocational knowledge is based largely on these frequent collocation types although the less dominant type is also included.

5. Discussion

The analysis of the COBN revealed that the majority of them are nouns (N = 99) which is accounted for 82.5%. The adjective (N = 15) comes the second, accounted for 12.5%. The smallest number of the keywords is verb (N = 6), constituting 5%. Unsurprisingly, the word 'pandemic' has the highest keyness value (8807.01) of them all; this could be due to the circulation of the COVID-19 outbreak on global news providers including BBC.com which is the primary source of language data during the period of the corpus construction. Evidently, the abounding number of contents words such as nouns, adjectives, and verbs in the keyword list in the COBN correspond to those found in a number of studies carried out by several scholars including Coxhead's (2000) Academic Word List, Chung's and Nation's (2004) technical vocabulary, Yang's (2012) keywords in hotel brochure text, Yang's (2015) nursing academic word list, Trinant's and Kijpoonphol's (2021) collocations of keywords in tourism texts, Suraprajit's (2022) collocations of logistics magazines, and Kadirbekova's **(**2023) collocations of keywords in ICT.

When considering the collocations according to the set framework adapted from Benson *et al*. (2010), it was disclosed that most collocations of keywords extracted from the COBN are under the set framework. The combination types of the lexical collocations encompass Noun + Noun (N = 152, 53.90%), Adjective + Noun (N = 77, 27.30%), Noun + Verb (N = 34, 12.05%), Verb + Noun (N = 11, 3.90%), and Verb + Adverb (N = 3, 1.06%). This high number and percentage of the collocation types denote that collocations under the adopted framework are dominant combination types. Specifically, the prominent number of Noun + Noun and Adjective + Noun combinations is similar to the study carried out by Biber *et al.* (2011) in which complex noun phrases were found to be dominant features in research articles.

Apart from the revelation of combination types, an examination of the examples of collocation pairs also yields useful insights into the grammatical properties of collocation pairs. For an illustration, there are some verbs collocates of noun nodes which are used in a passive form such as restrictions (were) lifted, bills (were) set, and company (was) founded. This could potentially provide pedagogical implications in the sense that this language feature should be recognised and highlighted to the students by English teachers.

6. Conclusion

The application of the corpus tool in identifying collocations from instructional materials such as textbooks, articles, news, and related documents could be of great benefit for ESP learners such as business English students as presented in this study, as well as in EAP and ESL/EFL contexts at large. By extracting and classifying collocations from instructional materials used in each course of a particular academic as well as making a list of those collocations, should make it easier and more focused for teachers in designing lessons and teaching materials as well as for learners in noticing those collocations. What is equally important is that corpus-based lessons should benefit learners in raising learners' awareness of collocations since they are significantly important in developing their natural and native-like fluency.

The study also suggests several recommendations for future research. Firstly, it encourages extending the focus beyond identifying lexical collocations to include the study of grammatical collocations. This could involve identifying grammatical collocations or comparing both types of collocations. Secondly, considering that the study focuses on business English, future research could explore collocations in various academic fields to enhance understanding and teaching across disciplines. Lastly, the study relied on online business news, and future research could diversify by using other authentic materials like textbooks, magazines, research articles, and business journals to provide a broader perspective on business collocations.

Conflict of Interest Statement
The authors declare no conflicts of interest.

About the Author(s)
Krittat Sukman (PhD) is a lecturer of English at the Faculty of Liberal Arts, North Bangkok University, Thailand. Prior to his PhD in English Language Teaching from Thammasat University, Thailand, he obtained MA in English Language Teaching and BA in Business English from Assumption University of Thailand. Recently, he was awarded an MBA from University of Northampton, United Kingdom. His research interests include corpus linguistics, vocabulary teaching, and Global Englishes.

References

Ackermann, K. & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, 12(4), pp. 235-247.

Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software/antconc/

Bennett, G. R. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. Ann Arbor, MI: Michigan University Press.

Benson, M., Benson, E. & Ilson, R. (2010). *The BBI combinatory dictionary of English: Your guide to collocations and grammar* (3rd ed.). Amsterdam: John Benjamins.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in l2 writing development? *TESOL Quarterly*, 45(1), 5-35.

Chen, L. (2017). Corpus-aided Business English Collocation Pedagogy: An Empirical Study in Chinese EFL Learners. *English Language Teaching*, 10(9), 181-197

Chung, T.M. & Nation, P. (2004). Identifying technical vocabulary. System, 32, 251-263

Coxhead, A. (2000). A new academic wordlist. *TESOL Quarterly*, 34(2), pp. 213-238.

Jones, S., & Sinclair J. McH. (1974). *English lexical collocations.* Cahiers de Lexicologie, 24, 15-61.

Kadirbekova, D. (2023). Corpus-based approach in learning collocations in information-communication technology (ICT). *Oriental Renaissance*, 3(2), 432-446.

Khamphairoh T, Tangpijaikul M (2012) Collocations of keywords found in insurance research articles: a corpus-based analysis. *Humanities Journal*, 19(2), 166–88.

Kiss, I. & Horváth, J. (2015). *Sheltered Beaches: A Tourism Collocation Approach to CLIL Vocabulary Teaching*, In Lehmann, M., Lugossy, R. & Horváth, J. (Eds.) Empirical Studies in English Applied Linguistics, Lingua Franca Csoport Pécs (pp. 166-178).

Henriksen, B. (2013). *Research on L2 learners' collocational competence and development – A progress report*. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), Vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis (pp. 29-56).

Hill, J. (2000). *Revisiting priorities: From grammatical failure to collocational success*. In M. Lewis (Ed.), Teaching collocation: Further development in the lexical approach (pp. 47-69). London: Commercial Colour Press Plc.

Hou, H. (2014). Teaching Specialized Vocabulary by Integrating a Corpus-Based Approach: Implications for ESP Course Design at the University Level. *English Language Teaching*, 7(5), 26-37

Lewis, M. (Ed.). (2000). *Teaching collocation: Further development in the lexical approach*. Oxford: Oxford University Press.

McCarthy, M. & O'Dell, F. (2005). English Collocation in Use. In Barnbrook, Mason, O. & Krishnamurthy (eds) Collocation applications and implications, *Palgrave Macmillan*, 2013, p. 120.

Menon, S., & Mukundan, J. (2012). Collocations of high frequency noun keywords in prescribed science textbooks. *International Education Studies*, 5(6), 149-160.

Molavi, A., Koosha, M., & Hosseini, H. (2014). A Comparative Corpus–based Analysis of Lexical Collocations Used in EFL Textbooks, *Latin American Journal of Content and Language Integrated Learning*, 7(1), pp. 66-81.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nelson, M. A. (2000). *Corpus-based study of business English and business English teaching materials*. (Electronic version). Unpublished PhD Thesis. Manchester: University of Manchester.

Newman, A. (1988). The Contrastive Analysis of Hebrew and English Dress and Cooking Collocations: Some Linguistic and Pedagogic Parameters. *Applied Linguistics*, 9(3), 293-305.

Önder Özdemir, N. (2014). Using corpus data to teach collocations in medical English, *Journal of Second Language Teaching and Research*, 3(1), 37-52.

Parkinson, J. (2015). Noun-noun collocations in learner writing. *Journal of English for Academic Purposes*, 20, 103-113.

Scott, M. (2012). WordSmith Tools (Version 6). [Computer Software]. Liverpool, UK: Lexical Analysis Software.

Suraprajit, P. (2022). English words and collocations found on logistics magazines: a corpus-based study. *Advances in Language and Literacy Studies*, 13(1), 41-18.

Tabak, M. & Takač, V.P. (2023). Relationship between collocational competence and collocation learning strategies in an English for specific purposes context. *International Journal of TESOL Studies*, 5(2), 113-131.

Trinant, K. & Kijpoonphol, B. (2021). Lexical collocations in a sample corpus of tourism research articles (SCTRA). *NKRAFA Journal of Humanities and Social Sciences*, 9, pp. 94–108.

Trinant, K. & Yodkamlue, B. (2019). Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA), 13(1), pp. 45-72

West, M. (1953). *A general service list of English words*. London: Longman.