CORPUS-BASED ANALYSIS OF COLLOCATION TYPES IN ONLINE NEWS ARTICLE ON THE VINA CASE FROM RADAR CIREBON WEBSITE

Mochamad Tabah

Program Studi Sastra Inggris, Institut Prima Bangsa Cirebon

mochamadtabah@gmail.com

ABSTRAK

Penelitian ini bertujuan menganalisis kolokasi dalam artikel berita daring mengenai kasus Vina di situs radarcirebon.disway.id. Data dikumpulkan dengan kata kunci "kasus vina" dan melibatkan 197 artikel yang disimpan dalam file .txt dan dimasukkan pada korpus. Menggunakan metode campuran, penelitian ini menggabungkan analisis kuantitatif dengan AntConc dan analisis kualitatif. Dari 5.460 kata yang dihasilkan, 100 kata teratas dipilih dan direduksi menjadi 57 kata konten setelah mengecualikan kata fungsi. Kata-kata tersebut dikategorikan berdasarkan bagian ujaran, dan kolokasi ditemukan berdasarkan theory Michael McCarthy dan Felicity O'Dell, yang mencakup enam jenis kolokasi, yaitu Adjektiva + Nomina, Nomina + Nomina, Verba + Nomina atau Nomina + Verba, Verba + Preposisi/Ekspresi, Verba + Adverbia, dan Adverbia + Adjektiva. Hasil analisis menunjukkan dua jenis kolokasi dominan: Verb + Noun (frekuensi 51) yang menggambarkan proses hukum dan Noun + Noun (frekuensi 2.067) yang menunjukkan elemen-elemen penting seperti pelaku, korban, entitas hukum, dan lokasi. Kolokasi yang ditemukan menunjukkan bagaimana media membingkai kasus Vina dengan menekankan aspek hukum dan prosedural. Media juga mengaitkan aktor-aktor penting seperti pelaku, korban, dan aparat penegak hukum dalam struktur narasi yang berulang dan bermakna, sehingga membentuk pemahaman publik yang terarah.

Kata kunci: *kolokasi, kasus vina, analisis korpus, berita daring.*

ABSTRACT

This study aims to analyze the collocation in online news articles about the Vina case on the radarcirebon.disway.id website. The data was collected with the keyword "kasus vina" and involved 197 articles stored in .txt files and included in the corpus. Using a mixed method, this study combines quantitative analysis with AntConc and qualitative analysis. From the 5,460 words generated, the top 100 words were selected and reduced to 57 words of content after excluding the word function. The words are categorized by parts of speech, and colocation is found based on the theory of (Michael McCarthy & Felicity O'Dell, 2005), which includes six types of collocation, namely Adjectives + Nouns, Nouns + Nouns, Verbs + Nouns or Nouns + Verb, Verbs + Prepositions/Expressions, Verbs + Adverbs, and Adverbs + Adjectives. The results of the analysis showed two dominant types of collocation: Verb + Noun (frequency 51), which describes the legal process, and Noun + Noun (frequency 2.067), which indicates important elements such as the perpetrator, victim, legal entity, and location. The collocation found shows how the media frames the Vina case by emphasizing legal and procedural aspects. The media also links important actors such as perpetrators, victims, and law enforcement officials in a repetitive and meaningful narrative structure, thus shaping a directed public understanding.

Keywords: collocation, vina case, corpus analysis, online news.

INTRODUCTION

Communication is a fundamental process where individuals or groups share information and ideas, conveying ideas, emotions, and desires through human interaction alone. In the media context, communication is crucial in influencing public opinion. News articles and online journalism are primary vehicles of communication, serving as the main sources through which information is disseminated to the public. This occurs because the media plays a role in providing information related to phenomena or events that occur in the lives of the general public (Faizal et al., 2018). Moreover, another perspective states that the media must deliver news quickly, accurately, and in a timely manner about events happening in society; additionally, news itself represents current and significant events (Denny Nathaniel Chandra et al., 2019). With the rise of digital platforms, online journalism has gained significant prominence. This evolution has turned online news outlets into influential players in shaping how news events. In a related linguistic context, collocation is an important concept that examines how words commonly occur together in language.

Along with this shift, the study of linguistic patterns such as collocation becomes increasingly relevant in understanding how information is framed and conveyed. Collocation is one of the important concepts in the study of linguistics, and it deals with how words are used together in a language. In the field of linguistics, colocation refers to pairs of words that often appear side by side in a natural context. In line with the definition given by (Michael McCarthy & Felicity O'Dell, 2005), a collocation is a pair or group of words that are often used together. Thao et al. (2021), as quoted by (Laoli et al., 2021), emphasized that collocation is not just a relationship of meaning but a relationship of the probability of the occurrence of certain words in one context. In other words, collocation indicates the tendency of two or more words to appear close together in actual use of language. There are six types of collocation according to (Michael McCarthy & Felicity O'Dell, 2005), namely Adjectives + Nouns, Nouns + Nouns, Verbs + Nouns or Nouns + Verb, Verbs + Prepositions/Expressions, Verbs + Adverbs, and Adverbs + Adjectives. These six types of collocation reflect the way words naturally pair with each other in language and show lexical patterns commonly used by native speakers.

Learning collocation is important in language mastery because it helps the speaker or writer use the language more naturally and effectively. (Michael McCarthy & Felicity O'Dell, 2005) states that collocation provides the most natural way to express things. In addition, collocation also enriches language expression by providing a more expressive or appropriate alternative. Collocation also contributes to the improvement of the writing style. By using the right collocation, the writing becomes more varied, structured, and easy to read. Therefore, an understanding of collocation is essential in the effective use of language, both in oral and written communication, mainly when used in academic or journalistic contexts.

To explore collocational patterns more systematically, corpus analysis offers a powerful method for studying language patterns and textual structures within large bodies of text. Corpus analysis is a highly effective method for studying language patterns and

textual structures within large bodies of text. Corpus linguistics is based on collections (bodies) of authentic texts stored on computers and analyzed using software designed for corpus analysis (Gena Bennet, 2010). Anatol Stefanowitsch (2020), stated that corpora used in corpus linguistics research consist of authentic language use, representing linguistic diversity in studies, and including a large amount of data. The use of corpus tools, such as AntConc, allows researchers to quantitatively analyze the occurrence of key terms and qualitatively assess how these terms are used within different contexts.

One of previous studies have been conducted to identify the types of collocation in different types of texts. Laoli et al. (2021), in his research entitled "Analysis on the Collocation Types Found in Novel Percy Jackson and The Olympians 'The Battle of Labyrinth'," found that of the six types of collocation according to McCarthy, only four appear in the novel, namely verb + adverb (75%), noun + noun (15%), noun + preposition (5%), and verb + noun (5%). Meanwhile, the adjective + noun and adverb + adjective types are not found at all. These findings show the dominance of action expressions and descriptions in building adventure narratives in literary texts.

Cynthia Veronika (2020), in her research "The Structure of Collocation and Its Translation Structures in Subtitles of Kurzgesagt YouTube Channel", found that the adjective + noun colocation type dominated with 133 out of 232 collocations analyzed. Meanwhile, verb + noun appear 45 times or about 19.40%. These results reflect the importance of using descriptive structures in the delivery of scientific information that is visual and narrative in nature, as well as how collocation affects readability and naturalness in translated texts.

SITI KHOLIJAH RAY (2022), in her thesis entitled "An Analysis of Grammatical Collocation in the Jakarta Post Online Newspaper", analyzed five types of grammatical collocation found in online news articles. These types include noun + preposition (4 items), noun + to-infinitive (16 items), noun + that-clause (6 items), preposition + noun (88 items), and adjective + preposition (3 items). Of the total 117 collocations found, the preposition + noun type is the most dominant. However, this study did not find the existence of three other types of grammatical collocation, namely, adjective + to-infinitive, adjective + that-clause, and verb patterns. These results show that in news texts, grammatical collocation has an important role in conveying information concisely and formally.

The manner in which journalists select words, construct narratives, and present the case details can significantly impact how the public understands and reacts to these events. The "Kasus Vina," a widely discussed legal case involving Vina, a victim of a violent crime in Cirebon. This case is an example of such an event that captured significant attention in Indonesian media. This case has been extensively covered in local news, especially in Radar Cirebon, where it became a focal point in discussions about crime, justice, and local issues. To guide this analysis, the study seeks to answer two primary questions:

1. What are the most frequent content words in the corpus?

2. What types of collocations are found in the corpus, and how frequently do they occur throughout the discourse?

METHOD

This study employed a mixed-methods approach combining quantitative and qualitative descriptive methods. According to John W. Creswell & J. David Creswell (2014), this approach involves collecting, analyzing, and integrating both types of data to gain a comprehensive understanding. Quantitative analysis is conducted using AntConc software to examine word frequencies and concordances. Qualitative analysis is then used to interpret the contextual meaning behind the data.

To determine the data, especially to answer the first research question in this study, all online news articles collected from radarcirebon.disway.id are processed using AntConc software. The table presents the 100 most frequently occurring words in the corpus, which serve as the basis for identifying the most prominent vocabulary used in the coverage of the Vina Case. To answer the second research question, concordance analysis is used to explore word usage in context.

The object of this research is an online news article from the radarcirebon.disway.id website that discusses Vina's case. The article search process was carried out using the keyword "Kasus Vina" to capture initial data. However, not all articles that contain these keywords are immediately used as the object of analysis. Only articles that are genuinely related to the Vina case are included in the corpus. Both in terms of information, development, and the context of the news selected as part of the research corpus.

The data in this research is written text in the form of online news articles published by Radar Cirebon. A total of 197 news articles were collected from the website. These articles were published from July 11, 2023, to February 7, 2025. The articles cover various aspects of the case, including updates on police investigations, legal proceedings, public reactions, and commentary from relevant authorities or community members. This File is then processed using AntConc, a corpus analysis tool, to extract word frequency lists, collocations, and concordance lines.

The data collection techniques used in this study follow a systematic approach. First, the researcher conducted a keyword search using the phrase "Kasus Vina" on the Radar Cirebon website to ensure that only articles directly related to the case were retrieved. From the search results, the researcher then conducted manual selection and filtering by selecting articles that specifically focused on Vina's case, as well as excluding news or other articles that only mentioned the case in passing. Next, the full text of each selected article is copied and pasted into a plain text (.txt) file individually. Non-textual elements such as images, ad text, and irrelevant news are excluded from the corpus. These files are then compiled into a single folder for further processing. A total of 197 articles that have been collected are combined into one main text file or organized in a format that AntConc can read. The File serves as the main corpus in the process of linguistic analysis. The following are the steps that the analysis uses for the data:

1. The researcher used AntConc Software to analyze the online news corpus related to the Vina case.
2. The researcher enters the data by selecting the File> Open Directory menu > Choose Corpora Folder > Open, until the File appears in the Corpus Files column.
3. The researcher accesses the Word List tab and presses the Start button to display a list of the most frequently appearing words.
4. The researcher first extracted the top 100 most frequently occurring words in the corpus.
5. After that, the researcher analyzed standard linguistic features, namely function words and dominant content words.
6. To filter out function words, researchers imported a list of stopwords from the https://www.ranks.nl/stopwords/indonesian site into AntConc.
7. A function word list is applied in AntConc to exclude words that do not convey the content.
8. The filtering results are shown in the table, which contains the most frequent content words.
9. After reducing the list of words based on the frequency of occurrence, the researcher grouped the words into categories of part of speech (word classes), such as nouns, verbs, adjectives, adverbs, and prepositions
10. Based on these groupings, the researcher then analyzed collocation with reference to the colocation theory of Michael McCarthy & Felicity O'Dell (2005), which includes patterns such as Adjective + Noun, Noun + Noun, Verb + Noun or Noun + Verb, Verb + Preposition/Expression, Verb + Adverb, and Adverb + Adjective.
11. The researcher presents examples of colocation in the form of tables to illustrate the pattern of word relationships in the context of discourse.
12. Finally, the findings are described narratively to provide a deeper interpretation of how the Vina case is represented in discourse through the patterns of collocation that emerge.

RESULT AND DISCUSSION

In this section, the researcher will describe the results obtained after conducting the analysis. From a total of 197 online news articles about the Vina case that were input into the corpus tool, a total of 5,460 words appeared in the Word List tab. From this list, the top 100 most frequently occurring words were selected for further analysis. These words are presented in Table 1 and serve as the foundation for identifying dominant lexical patterns and the overall thematic focus of the discourse.

Table 1. The 100 most common words

Rank	Word	Freq	Rank	Word	Freq
1	*yang*	2223	51	*bisa*	241
2	*dan*	1652	52	*sidang*	238
3	*di*	1432	53	*pelaku*	235
4	*vina*	1360	54	*adalah*	233
5	*kasus*	1316	54	*keterangan*	233

6	cirebon	1202	56	pihak	227
7	ini	1035	57	satu	223
8	itu	922	58	keluarga	220
9	terpidana	761	59	atau	209
10	tidak	748	60	rudiana	206
11	dari	673	61	hakim	205
12	dengan	654	62	belum	202
13	dalam	653	63	kata	201
14	tersebut	632	63	terjadi	201
15	ada	591	65	salah	198
16	untuk	525	66	dpo	196
17	juga	476	67	baru	193
18	eky	468	67	polri	193
19	oleh	465	69	melakukan	192
20	pada	465	69	telah	192
21	hukum	451	71	menjadi	189
22	sudah	421	72	polisi	188
23	tahun	419	73	sebagai	183
24	saksi	413	74	terkait	180
25	ke	406	75	hari	179
26	akan	402	76	dilakukan	171
27	saya	402	77	setiawan	170
28	bahwa	401	78	kejadian	169
29	pk	372	79	kita	164
30	para	369	80	nama	163
31	kami	355	81	bukti	158
32	polda	351	81	titin	158
33	mereka	338	83	pemeriksaan	157
34	saka	333	84	setelah	155
35	pembunuhan	322	85	lalu	152
36	dia	318	86	dirinya	150
37	pegi	318	87	namun	149
38	saat	317	88	kemudian	147
39	orang	313	88	lebih	147
40	kuasa	309	90	agustus	143
41	kepada	305	91	reza	142
42	karena	296	92	atas	139
43	tatal	288	92	bukan	139
44	kota	286	92	lagi	139
45	kembali	280	92	terhadap	139
46	tim	276	96	informasi	138
47	menurut	269	96	korban	138
48	masih	265	98	lokasi	137

| 49 | jabar | 258 | 99 | iptu | 135 |
| 50 | sudirman | 247 | 100 | persidangan | 134 |

The results are presented in Table 1, which includes 100 words that frequently appear in the corpus. The word "yang" is the number 1 word that often appears in the corpus with a frequency of 2223, followed by the word "dan" with a frequency of 1652 in the second place, and "di" in the third place with a frequency of 206. This is the result of a corpus that has not undergone reduction.

Furthermore, to analyze words that are more conceptually meaningful, the researcher reduces functional words. This reduction process is carried out by importing a list of stopwords from the ranks.nl website, which contains words that do not carry contextual meaning in the discourse. After that, this list of stopwords is extracted and applied in the AntConc Software to filter out those functional words.

The results of this reduction process can be seen in Table 2, which shows words that are more conceptually meaningful after the exclusion of functional words. This table provides a clearer picture of the words that truly describe the important elements of the discourse.

Table 2. Frequent Content Words Following Stopword Reduction

Rank	Word	Freq	Rank	Word	Freq
1	vina	1360	31	hakim	205
2	kasus	1316	32	kata	201
3	cirebon	1202	33	terjadi	201
4	terpidana	761	34	salah	198
5	untuk	525	35	dpo	196
6	eky	468	36	baru	193
7	hukum	451	37	polri	193
8	tahun	419	38	melakukan	192
9	saksi	413	39	menjadi	189
10	pk	372	40	polisi	188
11	polda	351	41	terkait	180
12	saka	333	42	hari	179
13	pembunuhan	322	43	dilakukan	171
14	pegi	318	44	setiawan	170
15	orang	313	45	kejadian	169
16	kuasa	309	46	nama	163
17	tatal	288	47	bukti	158
18	kota	286	48	titin	158
19	kembali	280	49	pemeriksaan	157
20	tim	276	50	agustus	143
21	menurut	269	51	reza	142
22	jabar	258	52	atas	139
23	sudirman	247	53	informasi	138
24	sidang	238	54	korban	138

25	*pelaku*	235	**55**	*lokasi*	137
26	*keterangan*	233	**56**	*iptu*	135
27	*pihak*	227	**57**	*persidangan*	134
28	*satu*	223			
29	*keluarga*	220			
30	*rudiana*	206			

It can be seen in Table 2 that after a reduction to the functional word list, the remaining 100 words are 57. These words are the result of processing that removes words that do not carry substantial meaning in the discourse, such as conjunctions, prepositions, and other words that appear frequently but do not contribute significantly to the understanding of the main content of the corpus.

The result of this reduction focuses more on content words, which directly represent important actors, actions, or concepts in the discourse. The remaining words are made up of different classes of words, namely nouns, verbs, adverbs, adjectives, and prepositions. These five classes of words have a crucial role in conveying core information in the corpus and types of collocation according to the theory of Michael McCarthy and Felicity O'Dell.

Here are the results of the analysis of the remaining words, grouped by word class. The table below shows the division of these words:

Table 3. Distribution of Word Classes in the Corpus After Reduction

Noun	*kasus, cirebon, terpidana, hukum, tahun, saksi, pembunuhan, orang, kuasa, kota, tim, sidang, pelaku, keterangan, pihak, keluarga, hakim, kata, dpo, polri, polisi, hari, kejadian, nama, bukti, pemeriksaan, informasi, korban, lokasi, persidangan, vina, eky, saka, rudiana, sudirman, jabar, setiawan, titin, reza, iptu, pegi, tatal, atas*
Verb	*terjadi, melakukan, menjadi, dilakukan*
Adverb	*kembali*
Adjective	*salah, baru, terkait.*
Preposition	*untuk, menurut*

Noun Collocation Patterns in the Discourse

According to McCarthy and O'Dell (2008), there are six types of collocation, namely Adjectives and Nouns, Nouns and Verbs or Verbs and Nouns, Noun + Noun, Verbs and expressions with preposition, Verbs and Adverbs, Adverbs and Adjectives. In this study, two types of collocation are found, namely Verb and Noun, and Noun and Noun. From the results of the word class distribution, there are 10 verb + noun colocations with a total of 51 occurrences, and 24 noun + noun colocations with a total of 2067. The following will explain the types of collocations found:

A. Verb + Noun

Verb + noun collocation is a combination pattern between verbs and nouns that often appear together in a text or discourse. This collocation is important in discourse analysis because it can show the frequent relationship between a particular action or process (verb) and the object or phenomenon involved (noun). This collocation not only describes the actions performed, but also provides an idea of the roles and relationships between and related events or objects. For a more detailed look at examples of verb + noun collocations found in the corpus of news articles about the Vina case, see Table 4 below.

Table 4. Frequency of Collocation Pattern: Verb + Noun

NO	Collocation	Frequency
1	*Menjadi + Saksi*	10
2	*Menjadi + Korban*	9
3	*Melakukan + Pemeriksaan*	7
4	*Terjadi + Pembunuhan*	5
5	*Melakukan + Pembunuhan*	5
6	*Melakukan + Tindakan*	5
7	*Dilakukan + Pemeriksaan*	4
8	*Terjadi + Kasus*	3
9	*Menjadi + Terpidana*	2
10	*Menjadi + Pelaku*	1
Total		51

Based on the analysis of verb + noun colocation in the news corpus of Vina's case, ten major colocation pairs were found with a total of 51 occurrences. The most dominant collocations were "Menjadi + Saksi" (10 times) and "Menjadi + Korban" (9 times), which showed the important role of witness and victim in discourse. Collocations such as "Melakukan + Pemeriksaan", "Melakukan + Pembunuhan", and "Melakukan + Tindakan" reflect the intensity of legal action, while "Terjadi + Pembunuhan" and "Terjadi + Kasus" describe events as facts. Passive forms such as "Dilakukan + Pemeriksaan" give a formal impression, and other collocations such as "Menjadi + Terpidana" and"Menjadi + Pelaku", affirm the process of legal labeling of actors. Furthermore, a more in-depth analysis will be conducted for some examples of verb + noun collocation that appear, to understand how this collocation forms a narrative in the reporting of Vina's case.

Table 5. Example of Collocation Pattern: Verb + Noun from discourse

Name of Articles	Collocation
Komentar Pegi Setiawan Hadir di PN Cirebon Dukung 6 Terpidana Kasus Vina **– September 04, 2024**	*Tampak hadir juga teman-teman Pegi Setiawan sesama buruh bangunan yang* ***menjadi saksi*** *pada sidang praperadilan Pegi di PN Bandung di antaranya adalah Bondol.*

Polisi Sudah Periksa 68 Saksi Kasus Vina, Berkas Perkara Pegi Setiawan Ditarget Rampung Pekan Depan **– June 12, 2024**	*Sejumlah saksi akan dipanggil oleh penyidik Ditreskrimum Polda Jabar untuk **dilakukan pemeriksaan**.*
Sekrup Ada Dagingnya Jadi Novum, Kasus Vina Cirebon Pembunuhan atau Kecelakaan Ini Keyakinan Keluarga **- July 18, 2024**	*Salah satu faktor yang membuat keluarga Vina yakin bahwa telah **terjadi pembunuhan**, lanjut Reza, adalah hasil visum terhadap jasad kedua korban.*

The analysis of verb + noun collocations within various articles pertaining to the Vina case reveals a consistent linguistic pattern that encapsulates the legal events and actions surrounding the investigation. For example, in the news article *"Komentar Pegi Setiawan Hadir di PN Cirebon Dukung 6 Saksi Kasus Vina"* (September 04, 2024), the collocation *"Menjadi + Witness"* shows the important role of witnesses in the legal process, especially those present in the pretrial hearing. The use of this collocation emphasizes how witnesses are an important part of compiling legal evidence and determining the course of the trial. In the following news article, "*Polisi Sudah Periksa 68 Saksi Kasus Vina, Berkas Perkara Pegi Setiawan Ditarget Rampung Pekan Depan*" (June 12, 2024), the collocation "*Dilakukan + Pemeriksaan"* reflects the intensity of the investigation conducted by law enforcement, by showing that the examination of witnesses in the Vina case is still being carried out to reveal the truth. This highlights the efforts of the police in digging deeper information from witnesses.

Meanwhile, in the news article titled "*Sekrup Ada Dagingnya Jadi Novum, Kasus Vina Cirebon Pembunuhan atau Kecelakaan Ini Keyakinan Keluarga* " (July 18, 2024), the collocation *"Terjadi + Pembunuhan"* was used to underline the fact that the family believed that the event that happened to Vina was a murder, not an accident. This collocation serves to confirm that an event, in this case a murder, has occurred and is at the heart of public debate and investigation. In the end, the verb + noun collocation describe the legal process involving various actors, such as witnesses and the police, as well as the main events that became the focus of the reporting of the Vina case. The use of this collocation shows the importance of the elements of law, investigation, and events that occur as part of the narrative built in the media discourse.

A. Noun + Noun

Another type of collocation identified was the combination of a noun with another noun. Noun + noun collocation refers to noun pairs that often appear together in text, creating specific meanings or describing relationships between two different entities. This collocation is important because it can show specific ways in a discourse to describe interrelated concepts, actors, or objects. The noun + noun collocation help clarify the structure and context of the event being discussed. To see examples of noun + noun collocations found in the corpus of news articles on the Vina case, see Table 8 below, which contains the frequency of occurrence of noun + noun collocations.

Table 6. Frequency of Collocation Pattern: Noun + Noun

No	Collocation	Frequency
1	*Kasus + Vina*	708
2	*Kuasa + Hukum*	271
3	*Terpidana + Kasus*	247
4	*Polda + Jabar*	240
5	*Kasus + Pembunuhan*	175
6	*Kota + Cirebon*	161
7	*Pihak + Keluarga*	49
8	*Sidang + PK*	79
9	*Keluarga + Vina*	36
10	*Pelaku + Pembunuhan*	22
11	*Pelaku + Utama*	13
12	*Keterangan + Saksi*	11
13	*Tim + Hukum*	10
14	*Persidangan + Kasus*	7
15	*Kejadian + Pembunuhan*	6
16	*Kasus + Cirebon*	8
17	*Kejadian + Kasus*	4
18	*Pemeriksaan + Saksi*	4
19	*Hakim + Sidang*	4
20	*Keluarga + Korban*	4
21	*Nama + Pelaku*	3
22	*Korban + Pembunuhan*	3
23	*Terpidana + Pembunuhan*	1
24	*Pihak + Korban*	1
Total		2067

In the analysis of noun + noun colocation in the online news corpus regarding the Vina case, 24 types of colocations were found with a total frequency of 2067. The most frequently occurring collocation is "Kasus + Vina" with a frequency of 708, which indicates that the entire discourse in these articles is focused on cases involving Vina. Other fairly dominant collocations are "Kuasa + Hukum" (271) and "Terpidana + Kasus" (247), which reflect the focus on the role of attorneys and individuals involved as convicts in the judicial process.

The "Polda + Jabar" collocation (240) shows the involvement of police officers from the West Java region, who also have an important role in the investigation. "Kasus + Pembunuhan" (175) highlights the nature of the case, which is related to murder, a crucial element in the news. In addition, "Kota + Cirebon" (161) shows the importance of geographical location as part of identity and context in this news.

Several other collocations, such as "Pihak + Keluarga" (49), "Keluarga + Vina" (36), and "Keluarga + Korban" (4), reflect the importance of the role of the family in these

cases, both in the context of support and as an affected party. Collocations such as "Pelaku + Pembunuhan" (22), "Pelaku + Utama" (13), and "Nama + Pelaku" (3) indicate an attempt to identify and describe who is responsible in the case.

Other collocations, such as "Sidang + PK" (79), "Persidangan + Kasus" (7), "Keterangan + Saksi" (11), and "Kejadian + Kasus" (4) emphasize the importance of legal process, examination, and witness testimony in building the legal narrative of this case. Some other collocations, such as "Kejadian + Pembunuhan" (6) and "Kejadian + Kasus" (4), depict the main events that are the focal point of the investigation and evidence process.

Overall, these 24 types of collocation with a total frequency of 2067 times show how the structure of the discourse in the online news directs attention to various key elements in Vina's case, from the identity of the perpetrator and victim to the legal process and the location of the incident. To strengthen the understanding of this colocation function in shaping the focus of discourse, the following are some examples of *noun + noun* colocation that will be further analyzed.

Table 7. Example of Collocation Pattern: Noun + Noun from discourse

Name of Articles	Collocation
Krisna Murti Cs Akan Bawa Saka Tatal Mantan Terpidana Kasus Vina ke PN Kota Cirebon, Ini Tujuannya **– June 09, 2024**	*Upaya ini dilakukan untuk bahan peninjauan kembali terhadap Saka Tatal yang merupakan salah satu terpidana dalam **kasus Vina** yang saat ini sudah bebas.*
Elza Syarief Ancam Politisi yang Pansos Lewat Kasus Vina Yang Mau Nyalon Bahaya, Nih. **- July 01, 2024**	*Menurut dia, upaya dari para **kuasa hukum** terpidana kasus pembunuhan Vina dan Eky di Cirebon delapan tahun lalu hanya bisa mengupayakan PK alias peninjauan kembali.*
Saka Tatal Bikin Pengunjung Sidang PK 6 Terpidana Kasus Vina Tak Kuasa Menangis **- September 13, 2024**	*Saka Tatal, mantan terpidana kasus kematian Vina dan Eky, bersaksi di sidang lanjutan Peninjauan Kembali (PK) enam **terpidana kasus** tersebut.*

The analysis of noun + noun colocation in the media discourse regarding the Vina case reveals how the media constructs legal identity, social legitimacy, and the continuity of the case narrative through noun pairs. In the article *"Krisna Murti Cs Akan Bawa Saka Tatal Mantan Terpidana Kasus Vina ke PN Kota Cirebon, Ini Tujuannya "* (June 9, 2024), the collocation *"Kasus + Vina"* is used to mark legal cases that are the center of public attention. The use of this phrase serves as the main identification in the overall discourse, simplifying the complexity of events into a single term that serves as a narrative anchor and a symbol of tragedy that is still legally and emotionally relevant.

Furthermore, in the article *"Elza Syarief Ancam Politisi yang Pansos Lewat Kasus Vina Yang Mau Nyalon Bahaya, Nih"* (July 1, 2024), the collocation of *"kuasa + hukum"* appears to refer to institutional actors involved in legal defense efforts for the convict. Here, legal representatives are positioned not only as formal representations in legal

processes, but also as active parties in public and political discourse, even mentioned in the context of the tension between law and political imagery.

Meanwhile, the article *"Saka Tatal Bikin Pengunjung Sidang PK 6 Terpidana Kasus Vina Tak Kuasa Menangis"* (September 13, 2024) features the collocation of "*terpidana + kasus*" used to describe the legal status of individuals who have been sentenced for their involvement in the Vina case. The embedding of this phrase, even in former convicts such as Saka Tatal, shows that the legal label remains firmly embedded in social memory and media discourse, indicating how a person's identity can continue to be associated with the case even though the legal status has changed.

Overall, the use of noun + noun collocation illustrates how the media constructs a narrative structure that not only conveys legal information but also perpetuates social status, shapes public perception, and maintains the continuity of stories in the news. Through these lexical patterns, the media subtly reinforces certain ideologies and identities, leaving lasting impressions on public memory and discourse.

CONCLUSION

This study shows that in the online news article about the Vina case on Radar Cirebon website, there is a significant pattern of using collocation in shaping media narratives. Of the 197 articles analyzed with a corpus approach using AntConc, it was found that the two most dominant types of collocation were verb + noun and noun + noun. These findings confirm that the media not only conveys facts, but also builds a framework of social understanding through repetitive and meaningful word choices.

The verb + noun collocation appears in 10 combinations with a total frequency of 51 times. This collocation is closely related to legal and investigative processes involving various actors such as witnesses, victims, perpetrators, as well as the main events in the case. This pattern shows how the media emphasizes the procedural and legal aspects of the Vina case as an important element in reporting.

Meanwhile, the noun + noun collocation appears in 24 combinations with a much higher frequency, which is 2,067 times. This type of collocation dominates the narrative structure and reflects how the media links various important entities, such as perpetrators, victims, law enforcement officers, and the location of the incident, to the stages of the trial, in a single unity of meaning. This pattern strengthens the discourse that shapes public perception, maintains the continuity of the story, and emphasizes the social position of each actor in the case.

Overall, this study reveals that collocation plays an important role in the construction of media discourse, both in terms of legal information and in framing public opinion on criminal cases that are widely highlighted.

REFERENCES

Anatol Stefanowitsch. (2020). *COPUS Linguistics: A Guide to the Methodology*. Language Science Press.

Cynthia Veronika. (2020). *THE STRUCTURE OF COLLOCATION AND ITS TRANSLATION STRUCTURES IN SUBTITLES OF KURZGESAGT YOUTUBE CHANNEL*. https://doi.org/10.33479/klausa.v4i01

Denny Nathaniel Chandra, Gede Indrawan, & I Nyoman Sukajaya. (2019). KLASIFIKASI BERITA LOKAL RADAR MALANG MENGGUNAKAN METODE NAÏVE BAYES DENGAN FITUR N-GRAM. *Jurnal Ilmu Komputer Indonesia (JIKI)*, *4*(2). https://doi.org/10.23887/jik.v4i2.2772

Faizal, M., Faqih Abdillah, M., Aulia Sari IMS, D., Setiadi, W., Octavia, D., Suhendari, W., & Widiatmoko Soewardikoen, D. (2018). *PENGGUNAAN WEBSITE PORTAL BERITA SEBAGAI MEDIA INFORMASI UNTUK MAHASISWA* (Vol. 2, Issue 1). Online. https://doi.org/10.31598/bahasarupa.v2i1.217

Gena Bennet. (2010). *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. University of Michigan Press ELT. https://doi.org/10.3998/mpub.371534

John W. Creswell, & J. David Creswell. (2014). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches* (Fifth).

Laoli, H., Sihombing, A. R., & Silalahi, D. E. (2021). ANALYSIS ON THE COLLOCATION TYPES FOUND IN NOVEL PERCY JACKSON AND THE OLYMPIANS "THE BATTLE OF LABYRINTH." In *Multidisciplinary Peer Reviewed Journal ISSN* (Vol. 7).

Laurence Anthony. (2024). *AntConc* (Version 4.3.1). https://www.laurenceanthony.net/software/AntConc

Michael McCarthy, & Felicity O'Dell. (2005). *English Collocations in Use*.

Radar Cirebon. (n.d.). *https://radarcirebon.disway.id/search/kata/?c=kasus+vina*.

SITI KHOLIJAH RAY. (2022). *AN ANALYSIS OF GRAMMATICAL COLLOCATION IN THE JAKARTA POST ONLINE NEWSPAPER*. https://etd.uinsyahada.ac.id/8851/1/1720300129.pdf