

Women's Representation in the News of "Sexual Violence" in Online News Media

Shierly Amalia Aura Hayati¹, Abdul Muhaiminul Aziz², Agus Ridwan³

¹*State University of Surabaya, Surabaya, Indonesia*

²*National Cheng Kung University, Taiwan*

³*State University of Surabaya, Surabaya, Indonesia*

amaliashierly@gmail.com

ABSTRACT

This research is motivated by the high rate of sexual violence against women in Indonesia, which is often represented unevenly by the mass media, so that it tends to corner victims and reinforce gender stereotypes. The purpose of this study is to analyze and compare the representation of women in the reporting of sexual violence in two mainstream online news media in Indonesia, namely *Kompas.com* and *Tempo.co*, using a corpus linguistic approach. The method used is a *mixed method*, which combines quantitative analysis through word frequency, collocation, and concordance. Meanwhile, qualitative analysis through a critical discourse analysis approach. The research data consisted of 50 news articles related to sexual violence against women collected through observation techniques and document studies. The results of the study show that there is a significant difference in discourse strategies between the two media. *Kompas.com* tends to represent women directly as individual victims (*victim-centered*) by highlighting the identity and personal experiences of the victims. In contrast, *Tempo.co* emphasizes sexual violence as an issue-centered social, legal, and institutional issue, where women are often positioned as part of an institutional, policy, or legal process object rather than as subjects with personal experience. Despite the differences in the style of presenting the news, the two media still tend to place women in passive and subordinate positions that are limited to their status as victims in need of protection.

Keywords: Women, Sexual Violence, Corpus Linguistics, Discourse Studies

INTRODUCTION

Sexual violence is any sexual act or attempt that is carried out by force against a person, regardless of the relationship between the perpetrator and the victim and the situation in which it occurs (WHO, 2021). In Indonesia, sexual violence against women is still a social problem that continues to recur every year in Indonesia. Based on Komnas Perempuan data in 2024, sexual violence is the most reported form of violence, reaching 17,305 cases or around 26.94% of the total reports of violence against women. This figure shows that sexual violence is still a major issue that must receive serious attention from the public. However, the recorded figures are believed not to represent the true condition because many victims are reluctant to report due to social stigma and lack of protection (Arsa et al., 2022).

The high rate of sexual violence is not only a legal and social issue, but also related to how the case is represented in the public space, especially the mass media. Online news media is the main source of information for people in understanding cases of sexual violence because of the large number of reports about this issue in the mass media (Tenriawali, 2018). The delivery of news by the media greatly determines how the public views victims and perpetrators of sexual violence. However, in practice, the media often over-exposes the victim's side so that it has the potential to blame the victim, for example, by highlighting clothing, nighttime habits, or the assumption that the victim lures the perpetrator (Indainanto, 2020). This is in line with representation theory, which states that the media does not just photograph reality, but constructs meaning through words or images that they choose to highlight (Eriyanto, 2012).

Until now, the mass media in Indonesia have not fully succeeded in representing gender issues in a fair and balanced manner. Many news reports still reinforce traditional stereotypes that place women in subordinate positions, such as sexual objects or passive domestic roles, thus ignoring the complexity and diversity of women's experiences in society (Supiandi, 2025). Gender-insensitive reporting can reinforce stigma against victims and obscure the fact that sexual violence is a structural problem related to gender inequality. The media tends to produce dramatic texts as if normalizing the issue of violence against women, which causes society to consider this issue as commonplace, thereby reducing public awareness and concern for this issue (Indainanto, 2020).

Studies on how women are represented in news texts have been conducted by many researchers before. Puteri et al. (2020) used Sara Mills perspective to analyze news in Metrojambi.com and found that even though women are positioned as subjects, they are still portrayed as weak and marginal figures. Meanwhile, Rahayu and Agustin (2020), through the framing model of Gamson and Modigliani found that Tirto.id media has a positive tendency by framing sexual violence because of patriarchy and inviting people to fight against this injustice. On the other hand, Safira et al. (2021), who applied Roger Fowler's model to news analysis on Republika Online, found that the use of passive sentences and euphemisms often marginalizes female victims. Although there has been a lot of research on women's representation, most of them still focus on qualitative analysis of news content on a limited basis and has not taken advantage of a comprehensive data-driven approach.

This research is here to fill this gap by using a corpus linguistic approach in discourse studies, which allows a systematic and evidence-based analysis of women's representation of a large number of news texts. Corpus linguistics is a branch of science that analyzes language by using a large collection of texts (corpus) to identify linguistic patterns such as frequency, collocation, and concordance that are not easily captured through manual reading (Biber & Reppen, 2015). The use of the corpus method allows

researchers to uncover how media texts tend to repeat certain patterns in framing issues over a period of time (Ahmed & Matthes, 2016). With the help of corpus software, researchers can extract keywords and the context of word occurrence (concordance) to identify how women are constructed discursively (Arsa et al., 2022).

The object of study in this study is two mainstream online news media in Indonesia, namely Kompas.com and Tempo.co. The selection of online media is based on the fact that this media has become the main source of information for the community that allows quick and interactive access (Haryatmoko, 2007). Kompas.com was chosen because of its reputation as one of the "10 Best Online Mass Media in the Use of the Indonesian Language" by the Language Agency of the Ministry of Education and Culture, Research and Technology. Meanwhile, Tempo.co is known for its sharp and in-depth investigative coverage, and is one of the most trusted media in Indonesia according to the Reuters Institute's Digital News Report 2022. Comparing these two major media with different news presentation characteristics is essential to uncover discourse patterns that may reinforce gender bias or even support equality.

Representation in media discourse is not only shaped by individual word choices, but by repetitive patterns that form a particular structure of meaning. If the word *korban* 'victim' in relation to women is often associated with terms such as *lemah* 'weak' or *tak berdaya* 'helpless', then this indicates the existence of an ideological structure that indirectly reproduces certain stereotypes (McEnery & Hardie, 2012). Therefore, this study uses the framework of Critical Discourse Studies (KWK) which views the dialectical relationship between texts (discourse), discursive practices (text production and consumption), and power structures in society (Fairclough, 2013). Through a corpus-driven approach, this study seeks to uncover the ideological constructions hidden behind the word choices and narrative structures used by online news media Kompas.com and Tempo.co.

In addition, this research has strong relevance to the 5th *Sustainable Development Goals* (SDGs), namely Gender Equality. Fair representation of women in the media is very important because the media has great power in shaping social norms. According to UN Women (2022), if used responsibly, the media can advance gender equality and women's empowerment. On the contrary, biased reporting can strengthen patriarchal structures. Thus, the results of this study are expected to be a reference for efforts to improve journalistic practices that are more in favor of gender justice and encourage the formation of a more empathetic and inclusive media narrative towards victims of sexual violence.

METHOD

This study uses a *mixed-method* approach, which is an approach that combines quantitative and qualitative. A qualitative approach is used to describe the representation of women in news texts through critical discourse analysis and interpreting the meaning of the corpus findings. Meanwhile, a quantitative approach is used in analyzing corpus linguistics to measure frequency, collocation, and concordance in news texts. The research data was in the form of word frequency, collocation and concordance contained in news articles about sexual violence against women in online news media *Kompas.com* and *Tempo.co*. The data used consisted of article texts published in *Kompas.com* and *Tempo.co* in the period January to September 2025. In this study, 25 news articles from *Kompas.com* and 25 news articles from *Tempo.co* focused on sexual violence against women were collected. Once collected, all the data is formatted in the form (.docx) and then converted to the format (.txt) to be analyzed using the AntConc software.

The data collection technique of this study uses observation techniques and document studies. The observation technique was used to observe the text of news articles in online *media Kompas.com* and *Tempo.co* that focused on reporting sexual violence against women. Meanwhile, the document study technique was carried out by collecting and processing the text of news articles from online *Kompas.com* and *Tempo.co* media related to sexual violence against women. This technique functions to complement observations by utilizing written documents as a source of data (Sugiyono, 2010). The process includes observation and selection of relevant articles, copying text to *Word* documents, removing non-text elements (such as dates, advertisements, and links), compiling corpus based on media, converting files to formats (.txt), and analyzing using *AntConc* to obtain frequency, collocation, and concordance data related to women's representation.

This data analysis technique is carried out through the following steps: 1) Collecting the text of online news articles related to sexual violence against women from *Kompas.com* and *Tempo.co* media then compiling it into corpus data in text format (.txt). 2) Using *the AntConc* software to perform word frequency analysis. This step aims to find words that are often used to represent women in the reporting of sexual violence. 3) Analyze word frequency and word collocation data. Furthermore, corpus-based concordance analysis that often appears represents women in the reporting of sexual violence. The corpus that has been obtained is then presented in the form of a data analysis table. 4) Exploring the results with Critical Discourse Analysis (AWK) to interpret how words and collocation shape women's representation in reporting sexual violence.

RESULTS AND DISCUSSIONS

Frequency of Words in the News of Sexual Violence Against Women in Online News Media Kompas.com.

In the January-September 2025 period, there were 25 news articles collected with a total of 1,698 word types and a total of 7,409 word tokens. Based on the results of the corpus, some words often appear in the news of sexual violence against women in the following order of frequency.

Table 1. Word Frequency on Kompas.com

No.	Sequence Frequency	Quantity Frequency	Laughter
1	1	173	korban 'victims'
2	6	82	pelecehan 'harassment'
3	7	81	seksual 'sexual'
4	9	71	pelaku 'offender'
5	12	64	perempuan 'women'
6	17	52	dia 'he/she'
7	22	41	kekerasan 'violence'
8	23	40	kasus 'case'
9	26	36	saya 'I'
10	29	34	laki 'male'

Table 1 found 10 relevant words from the top 30 frequencies in reporting sexual violence against women. This list is the result of filtering from *AntConc's initial data*, which was previously dominated by task words such as *di* 'in', *yang* 'that', and *dan* 'and'.

From the results of the analysis, it shows that the words "victim, harassment, sexual, perpetrator, woman" are the words that appear most often in the news corpus analyzed. The word "victim" became the word with the highest frequency, which was 173 times. The high frequency of use of the word "victim" shows that this term is the main reference in reporting sexual violence against women. However, at the frequency analysis stage, the word "victim" cannot be confirmed to refer directly to women because it has not been accompanied by an explicit reference.

In addition, the word "victim" and the word "harassment" occupy the second and third most frequent positions, with 82 and 81 appearances, respectively. The high frequency of the word "harassmen" shows that the media more often frames sexual violence as an act of harassment.

The word "sexual" appears 81 times, almost equal to the word "harassment". The appearance of the word "sexual" generally coexists with the word *kekerasan* 'violence' or *pelecehan* "harassment." The word "sexual" serves as a marker of the type of violence in the news and reinforces the focus of the online news media on the aspect of sexual acts themselves.

Furthermore, the word "perpetrator" appears 71 times. The word "perpetrator" indicates the media's efforts to display the relationship between the perpetrator of the violence and the victim. However, in the reporting of sexual violence, often the identity and background of the perpetrator are not discussed in depth, so the focus of the narrative is more on displaying women as victims than perpetrators.

Lastly, the word "woman" appeared 64 times in the news corpus analyzed. The use of the word "woman" in the news generally refers to gender identity in cases of sexual violence.

Word Collocation in the News of Sexual Violence Against Women in Online News Media Kompas.com

Based on the 25 news articles analyzed, the *Total Number Of N-Gram Types* were obtained as many as 1,458 and *the Total Number of N-Gram Tokens* as many as 4,064. The appearance of the word "women" in this collocation analysis is differentiated based on its position, namely in the left position (*on left*) and in the right position (*on right*). The word "woman" in the left position (*on left*) indicates that women are the initial element which is then followed by certain information. Meanwhile, in the right position, *the* word Perempuan 'woman' becomes an element that is described or as an object in the sentence. This division is based on the presentation and collocation data in the following table.

Table 2. Collocation of the word "female" (*on left*) in *Kompas.com*

No.	Sequence Frequency	Quantity Frequency	N-Gram/Cluster
1	27	9	perempuan dan anak 'women and children'
2	185	4	perempuan berinisial 'women with initials'
3	186	4	perempuan dan perlindungan 'women and protection'
4	356	3	perempuan jadi korban 'women are victims'
5	357	3	perempuan korban 'women are victims'
6	1132	2	perempuan menjadi korban 'women become victims'

In the table data above, the word "woman" consistently appears in the left position and collocates with several words in the form of nouns and verbs. The highest frequency collocation is 'women and children', which appears 9 times. The collocation with the noun can be seen in the data of "women and children", which are used as subjects or objects in reporting sexual violence as a group that is often targeted.

Furthermore, the data on "women with initials" places women as the subject of identities disguised through initials. Other data, such as "women and protection" and P "female victims" represent women as subjects attached to the status of victims and parties associated with protection.

Meanwhile, collocation involves verbs, such as "women are victims" and "women become victims". The use of the verbs *jadi* 'to be' and *menjadi* 'being' shows that women are

No.	Sequence Frequency	Quantity Frequency	N-Gram/Cluster
1	53	7	seorang perempuan 'a women'
2	65	6	Komnas perempuan 'National Commission of women'
3	106	5	pemberdayaan perempuan 'empowerment women'
4	109	5	perlindungan perempuan 'protection women'
5	277	3	Jurnalis perempuan 'women journalist'
6	299	3	Kementerian Pemberdayaan Perempuan 'Ministry of Empowerment Women'
7	418	3	Unit Perlindungan Perempuan 'Protection-Unit Women'
8	641	2	Dinas Pemberdayaan perempuan 'Empowerment Office Women'
9	963	2	melecehkan perempuan 'harass women'

represented as parties who experience actions or events in the reporting of sexual violence.

Table 3. Collocation of *the word woman* (on right) in Kompas.com

In table 3, the word "woman" consistently appears in the right position and as a noun described by the previous word. The second-highest frequency collocation is "a woman," which appears 7 times. This collocation is a noun that places women as the sole subject in the reporting of sexual violence.

Furthermore, most of the data in Table 3 show the dominance of noun collocations related to the names of institutions and policies, such as "*Komnas Perempuan*", "*Women's Empowerment*", "*Women's Protection*", "*Ministry of Women's Empowerment*", "*Women's Protection Unit*", and "*Women's Empowerment Office*." The collage is classified as a noun because it refers to a noun in the form of the name of an institution, ministry, unit, or action that plays a role in the protection and handling of women.

In addition, there is a collocation that associates women with a profession or social role, namely "women journalists". This collocation is also a noun that functions as a unity of meaning and can occupy the position of the subject or object in a sentence.

Meanwhile, the collocation "*Harassing Women*" is a verb. The verb "Harass" is used to describe an action or situation experienced by a woman. This also shows women in the news as parties who experience acts of sexual violence.

Based on the results of the collocation analysis in tables 2 and 3, it can be interpreted that the reporting of sexual violence tends to portray women as victims who need protection, both individually and as a group. The representation of women as active or empowered subjects still seems limited, as the media focuses on news that emphasizes women's experiences of violence and the role of institutions in handling these cases.

Concordance of Women's Words in Online News Media Kompas.com

After frequency and collocation analysis, the next stage is concordance analysis to look more deeply at the use of the word "woman" in the reporting of sexual violence. The analysis used a single keyword, "women," focusing on the top 20 data because the word most explicitly refers to the main subject. By focusing the analysis of concordance on one keyword, this study can explore more deeply the variation in the context of the use of the word "woman" in news texts, so that the pattern of women's representation can be observed more easily. The results of the analysis of the concordance of the word "female" are shown in Figure 1 as follows.

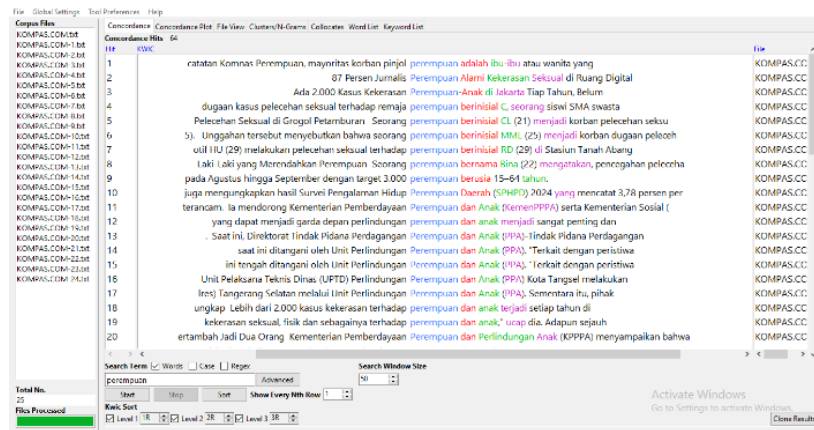


Figure 1. Concordance of the word woman in Kompas.com

Based on the concordance in figure 1, the word "woman" appears in various contexts, both as individual data sections, and related to protection institutions. In the initial data, women are represented as a socio-economically vulnerable group, for example in the quote "the majority of female loan victims are mothers or women". In addition, women also appear as victims of sexual violence in the digital space through the quote "87 percent of female journalists experience sexual violence in the digital space". Combining with children in statistics, such as "2,000 cases of violence against women and children in Jakarta every year", further confirms the position of women as a vulnerable group.

In data 4-7, women are represented as individual victims through the mention of initials, such as "adolescent girl with the initials C", "woman with the initials CL (21)", "woman with the initials MML (25)", and "woman with the initials RD (29)". This pattern shows efforts to protect identity, but at the same time places women especially as parties who

experience violence. However, there is a quote showing women as a more active subject, in the quote *"a woman ... say, prevention of harassment"*, which shows that women are not always positioned passively.

Furthermore, the institutional discourse seems dominant through the mention of *"Women and Children Protection Unit (PPA)"*, *"Women and Children Empowerment (KemenPPPA)"*, as well as the quote *"becoming the vanguard of the protection of women and children"* describing the state's concern and involvement in handling sexual violence, the way news like this indirectly forms the image of women as parties who are often placed as recipients of protection, compared as subjects who are active in sexual violence.

In addition, the use of quantitative data such as *"more than 2,000 cases of violence against women and children occur every year"* and the target of *"3,000 women aged 15-64"* also broadens the scale of the issue, but at the same time shifts women into a part of the statistics. In general, the news still tends to represent women as a vulnerable group and recipients of protection, although they occasionally appear as active subjects.

Frequency of Words in Sexual Violence News in Online Media Tempo.co

In the January-September 2025 period, there were 25 news articles collected with a total of 2.147 words types and a total of 10.842 words *tokens*. Based on the results of the corpus, there are words that often appear in the news of sexual violence against women in the following order of frequency.

Table 4. Word Frequency on Tempo.co

No .	Sequence Frequency	Quantity Frequency	Laughter
1	1	355	korban 'victims'

Table 4 (Continued)

No .	Sequence Frequency	Quantity Frequency	Laughter
2	5	150	seksual 'sexual'
3	7	110	kekerasan 'violence'
4	10	80	pelaku 'offender'
5	17	66	kasus 'case'
6	19	63	pelecehan 'harassment'
7	21	59	anak 'children'

8	27	48	perempuan 'women'
9	30	42	polisi 'police'
10	32	38	dia 'he/she'

In the table above, 10 relevant words were found from the top frequencies in reporting sexual violence against women in Tempo.co after going through *AntConc's* initial data screening and deletion of assignment words.

The results of the analysis showed that the words "victim", "sexual", "violence", "perpetrator", and "case" were the most dominant words that appeared. The word "victim" has the highest frequency, at 355 times, which shows that the main focus of the news is mostly directed at those who experience violence, although not all of them explicitly refer to women.

Furthermore, the word "sexual" appears 150 times and the word "violence" 110 times, which serves to emphasize the type of action that is reported as a serious problem. The word "perpetrator" appeared 80 times, lower than "victim", indicating greater media attention on the victim's position. Meanwhile, the word "case" appears 66 times and refers to an incident of sexual violence that is being processed or reported.

Collocation of Words About Women in Online News Media Tempo.co

Based on 25 news articles analyzed the *Total Number Of N-Gram Types* were obtained as many as 7,413 and *the Total Number of N-Gram Tokens* as many as 10,817. The appearance of the word "female" in this collocation analysis is differentiated based on its position, namely in the left position (*on left*) and in the right position (*on right*). The word "woman" in the left position (*on left*) indicates that women are the initial element which is then followed by certain information. Meanwhile, in the right position, *the* word "woman" becomes an element that is described or as an object in the sentence. This division is based on the presentation and collocation data in the following table.

Table 6. The Collocation of the Female Word (*on left*) in Tempo.co

No.	Sequence Frequency	Quantity Frequency	N-Gram/Cluster
1	200	5	perempan dan anak 'women and children'
2	346	4	perempuan dan perlindungan 'women and protection'
3	809	3	perempuan berusia ... 'women aged ...'

4	2152	2	perempuan usia 'women of ... age'
5	2154	2	Perempuan WNI 'Indonesian Woman'

Based on the data in the table, the word "women" appears in the left position (*on left*) which serves as the main noun to describe the word after it. The highest frequency collocation in this position is "women and children" which appear 5 times. This collocation is a noun that places women and children as a social group as a subject that is often referred to in the news of sexual violence.

The next collocation also has a relationship with the noun as shown in the data on "women and protection" which appears 4 times. This collocation shows women as subjects attached to efforts to form protection in sexual violence.

In addition, there is a collocation of "women aged ..." who appear 3 times and "women of ... age" 2 times. These two collocations show that the media often portrays women based on age range. By mentioning age, the media can still convey an overview of identity and maintain the privacy of women involved in sexual violence.

Meanwhile, the collocation of "Indonesian women" appeared 2 times. The collocation confirms the citizenship status of the women who are reported. By conveying citizenship status, the media helps readers understand that the events reported are within the scope of Indonesian law.

In the table 7, the word "woman" appears consistently appearing in the right position (*on the right*) as a noun described by the previous words. The highest frequency collocation is "girls" which appears 5 times. The collocation is a noun that functions as a subject or object in the reporting of sexual violence as a group that is often targeted.

No.	Sequence Frequency	Quantity Frequency	N-Gram/Cluster
1.	162	5	anak perempuan 'girls'
2.	292	4	kekerasan terhadap perempuan 'violence against women'
3.	340	4	pemberdayaan perempuan 'empowerment women'
4.	735	3	Menteri Pemberdayaan Perempuan 'Minister Women's empowerment'
5.	815	3	perlindungan perempuan 'women protection'
6.	1635	2	kekerasan pada perempuan 'violence on women'
7.	2245	2	remaja perempuan 'adolescent'

				Table
			girls'	
7.	8.	2381	2	seorang perempuan 'a women'
Collocation of the word woman (on right) in Tempo.co				

In addition, the collocation of "*violence against women*" and "*violence against women*" appeared 4 and 2 times, respectively. The two collocations show that women are objects or parties who are subjected to acts of violence. This strengthens the representation of women in the news as parties who experience the impact of sexual violence.

Furthermore, the collocation of "*women's empowerment*", "*women's protection*", and "*minister of women's empowerment*" shows the relevance of institutions and policies towards women in reporting sexual violence. The three collocations are nouns because they refer to the name of an institution or effort related to women's issues.

Meanwhile, the collocation of "*adolescent girls*" appeared 2 times, as well as the collocation of "*a woman*" which also appeared 2 times. These two collocations are nominal names that place women as the sole subjects in the reporting of sexual violence.

Concordance of Women's Words in Online News Media Tempo.co

After frequency and collocation analysis, the next stage is concordance analysis to look more deeply at the use of the word "woman" in the reporting of sexual violence. The analysis used a single keyword, "women," focusing on the top 20 data because the word most explicitly refers to the main subject. By focusing the analysis of concordance on one keyword, this study can explore more deeply the variation in the context of the use of the word "woman" in news texts, so that the pattern of women's representation

can be observed more easily. The results of the analysis of the concordance of the word "woman" are shown in figure 2 as follows.

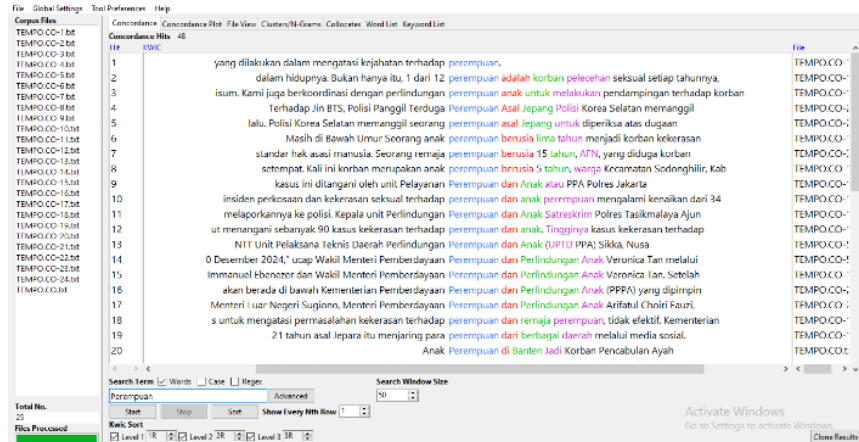


Figure 2. Concordance of the Word Women in Tempo.co

Based on the correlation analysis in table 8, the word "woman" in the reporting of sexual violence in *Tempo.co* appears in various contexts, such as victim status, forms of violence, mention of age and identity, and its relation to protection institutions. This shows that the representation of women is diverse and not singular in news texts.

In data 1 and 18, women are represented as the main targets of crime through the quote "... to address crimes against women" and "to address the problem of violence against women", which emphasizes protection efforts. In data 2, women are explicitly positioned as victims through the quantitative data "1 in 12 women are victims of sexual harassment every year", which confirms the high rate of sexual violence.

Meanwhile, data 3 places women with children in the quote "We are also coordinating with the protection of women and children to provide assistance to victims...", so that women are positioned as a vulnerable group. In data 4–8, women are shown as individual victims through common mentions such as "a Japanese woman", "a five-year-old girl", and "a 15-year-old girl". This way of presenting news protects the identity of the victim, but also limits the representation of women as the suffering party.

The consistency of the "women and children" pairs in data 9–17 such as "women and children's services", "sexual violence against women and children", "protection of women and children", and "Minister of Women and Child Empowerment" affirms women as a group that is considered vulnerable and in need of protection.

In data 19, the quote "capturing women from various regions through social media" represents women as targets of exploitation in the digital space. As for the 20 data, the statement "girls in Banten are victims of paternal abuse" shows the vulnerability of female victims in the family environment.

Comparison of Women's Representation in Reporting Sexual Violence in Online News Media Kompas.com and Tempo.co

Based on the results of the word frequency analysis, there is a difference in the occurrence of words between *Kompas.com* and *Tempo.co* in reporting sexual violence against women. The word "victim" is the word with the highest frequency in both media. However, the number of occurrences of the word "victim" in *Tempo.co* much higher, which is as much as 355 times compared to *Kompas.com* only 173 times. This shows that *Tempo.co* more often portrays the position of women as victims in the news. Meanwhile, on *Kompas.com* It also features women as victims, but in lower numbers, so the victim's narrative tends to be more limited to certain cases.

From the results of the comparison, the word "woman" also appeared in both media, with a frequency of 54 times in *Kompas.com* and 48 times in *Tempo.co*. This difference suggests that *Kompas.com* are relatively more likely to explicitly mention their gender identity in the news. Meanwhile, *Tempo.co* tend to emphasize the status of "victim" more than the direct mention of women's identities.

Although in the frequency analysis it has not directly shown that the word "victim" refers to women. This connection can be understood because the corpus in this study specifically focuses on reporting sexual violence against women. This was then strengthened through a collocation and concordance analysis that showed that the word "victim" often appears side by side with the word "woman".

Then, in the results of the analysis of the collocation of the two media, there is a clear difference in the way women are represented in the reporting of sexual violence. On *Kompas.com*, found collocations that explicitly place women as victims, as in the data "*Women become victims*", "*Female Victim*", and "*Women are victims*". The emergence of this collocation shows that *Kompas.com* directly associate women's identity with their status as victims of sexual violence. In other words, women are represented explicitly through their experiences of violence, so that women's position as victims appears explicitly in the structure of the language of the news.

On the contrary, in *Tempo.co* no collocation was found that directly stated "*female victim*" or "*woman becomes victim*". The collocations that appear in *Tempo.co* tend to be more general and structural, such as "*violence against women*", "*violence against women*", and "*girls*". This shows that *Tempo.co* emphasize violence as a social phenomenon or problem, rather than on affirming the status of women who are victims explicitly in the news.

This difference in collocation shows the difference in the way women are represented between the two media. In online news media, *Kompas.com* tend to build women's representation through a *victim-centered* approach, namely by explicitly prioritizing the

position and experience of the victim. Meanwhile, online news media Tempo.co more emphasizes *an issue-centered* approach, by placing sexual violence events as a social problem, so that the representation of women does not always appear as victims in the news.

In the consequent analysis, there is also a clear difference in the arrangement of news sentences between *Kompas.com* and *Tempo.co* in representing women in cases of sexual violence. This difference can be seen from the position of the word "woman" in the sentence and the context of the accompanying news.

In *Kompas.com*, the word "woman" tends to appear in sentences that describe the victim's condition, experience or identity directly. This can be seen in quotes such as "*the majority of female loan victims are mothers or women*" and "*87 percent of female journalists experience sexual violence in the digital space*". *Kompas.com* media highlights women's experiences as victims in various aspects of life, such as social, economic, and work. In addition, the mention of the victim's identity through initials and age, such as the quotes "*a teenage girl with the initials C*" and "*a woman with the initials CL (21)*", show that the media *Kompas.com* more often portrays women individually, but the representation is still limited to the violent events that befell them.

On the other hand, in *Tempo.co*, it tends to show more incidents of sexual violence and the process of handling it by official institutions in formal language, rather than exploring personal experiences women as victims. This can be seen in the quotes "*carried out in overcoming crimes against women*" and "*we also coordinate with the protection of women and children*". In the quote, the word woman is not always placed as the main subject in the news, but as part of the object of a policy, protection, or legal action. In addition, in the quote "*South Korean police summon a woman from Japan for questioning*", women more often appear as parties involved in legal proceedings than as women with experience of being victims.

These differences in sentence structure suggest that *Kompas.com* use more sentence patterns that focus on the victims and individual experiences of women. Meanwhile, *Tempo* tends to use sentence patterns that focus on events, policies, or institutional actions.

Thus, this study shows that the difference in women's representation between the two media is not only seen from the frequency and collocation, but also from the way news sentences are composed. This is because it affects the way readers understand the position and role of women in reporting sexual violence.

Women's Representation in Reporting Sexual Violence in Online Media Kompas.com

Based on the results of the analysis of frequency, collocation, and concordance, the representation of women in the reporting of sexual violence on the *Kompas* online news

media shows a strong tendency to position women as victims. This can be seen from the dominance of the appearance of the words "victim" and "woman" in the frequency analysis, which is then emphasized through the analysis of collocation such as "female victim", "woman becomes victim", and "woman and child". Meanwhile, the analysis of concordance shows that the word "woman" often appears in sentence structures that place her as a party to violence, either through the mention of identity with initials and age or through the incorporation of women in the category of vulnerable groups.

The findings of this study are in line with the research of Fitriana, et al. (2023). Fitriana et al.'s research on detik.com media found that women in reporting sexual violence tend to be represented as those who experience suffering and need protection. The study, especially in Kompas.com media, found that women still tend to be represented as victims and put them in subordinate positions. This shows that the results of this study strengthen the results of Fitriana et al.'s research.

Women's Representation in Reporting Sexual Violence in Online Media Tempo.co

Based on the results of the analysis of frequency, collocation, and concordance, the representation of women in the reporting of sexual violence in online news media *Tempo.co* shows a different tendency of Kompas.com. Although the word "victim" appears with very high frequency, the analysis of the collocation shows that *Tempo.co* does not display direct collocation such as "*Female Victim*" or "women become victims". On the other hand, women more often appear in the form of sentences that are general and structural, such as "*Violence against women*", "*Women's Protection*", and "women's empowerment". This shows that *Tempo.co* It does not emphasize women as explicit individual victims, but as a social category in the issue of sexual violence. Concordance analysis also shows that the word "woman" often appears in sentence structures related to policy, data, and legal processes, rather than in personal narratives as victims.

The findings of this study are in line with the research of Agphyra and Hamdani (2025). Agphyra and Hamdani's research found that women are more often represented as objects in the discourse of sexual violence, while women's personal experiences and voices as victims are rarely shown in depth. This study, especially in tempo.co, found that women are parties who need structural handling and protection, not as active subjects in news discourse.

CONCLUSIONS

Based on the results of research and discussion of women's representation in the reporting of sexual violence in online news media Kompas.com and Tempo.co using the study of corpus linguistic discourse, it can be concluded that both media still tend to

represent women as those who experience suffering and are in vulnerable positions. This can be seen from the dominance of the use of words such as *victim, harassment, violence, sexual, perpetrator* and collocation which consistently associate the word woman with victim status and the need for protection. This pattern shows that women are more often positioned as objects of events, rather than as active subjects in news discourse.

The results of the analysis of frequency, collocation, and concordance show that Kompas.com more prominently highlight women in the form of protection and institutions, such as through the mention of the Women and Children Protection Unit or the Ministry of Women's Empowerment. Meanwhile, Tempo.co tend to present news with a stronger emphasis on legal aspects and the process of handling cases, characterized by the high appearance of citations related to the law, apparatus, and investigation procedures. Despite the differences in reporting styles, both media places women in the position of passive subjects attached to their experiences of violence.

Overall, the findings of this study reinforce the view that online news media as the mainstream in Indonesia has not fully presented a balanced representation of women in the reporting of sexual violence. Women are still more often represented as victims in need of protection, while the representation of women as subjects who are actively speaking out and empowered is still very limited.

In addition, the results of this study show a difference with previous research, especially research by Rahayu and Agustin (2020) which found that the media Tirto.id framing sexual violence as a structural problem related to patriarchy and tending to place women as subjects involved in the narrative of resistance to injustice. In contrast to these findings, this study shows that Kompas.com and Tempo.co are more dominant using factual and institutional approaches, so that women are more often shown as news objects and policy recipients, rather than as active subjects. What needs to be noted is that the difference in the research object (*Tirto.id* with *Kompas.com* and *Tempo.co*) turns out to show a difference in the research results.

REFERENCES

- Ahmed, S., & Matthes, J. (2016). *Media Representation and Impact on the Lives of Women*.
- Arsa Ilmi Budiarti, et. al. (2022). Reflections on the Handling of Sexual Violence in Indonesia. *In Indonesia Judicial Research Society*.
- Barlow, M. (2004). *Collocate: A Corpus Analysis Tool*.
- Biber, D., & Reppen, R. (2015). *The Routledge Handbook of Corpus Linguistics*. Routledge.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (4th ed.)*. SAGE Publications.

- Eriyanto. (2012). *Framing Analysis: Construction, Ideology, and Media Politics*. Yogyakarta: LKiS Group.
- Fairclough, N. (2013). *Critical Discourse Analysis: The Critical Study of Language (2nd ed.)*. Routledge.
- Fitriana, et al (2025). Women's Imagery in Sexual Violence News: A Linguistic Analysis of the Corpus. *Pendas: Scientific Journal of Basic Education*.
- Haryatmoko. (2007). *Communication Ethics: Media Manipulation, Violence, and Pornography*. Yogyakarta: Kanisius.
- Indainanto, Y. I. (2020). Normalization of Women's Sexual Violence in Online Media. *Journal of Communication*, 14(2), 105-118. <https://doi.org/10.21107/ilkom.v14i2.6806>
- Komnas Perempuan. (2024) *Catahu 2023: Annual Record of Violence against Women in 2022 (Violence against Women in the Public and State Spheres: Lack of Protection and Recovery)*. <https://komnasperempuan.go.id/download-file/1316>
- Kridalaksana, H. (2008). *Linguistic Dictionary (Fourth Edition)*. Jakarta: Gramedia Pustaka Utama.
- Kusuma, A., Azharani, R. P., Sitiyani, I., & Thusyifa, N. (2024). Framing Online Media: Sexual Violence Against Women in CNN Indonesia News. *Journal of Tambusai Education*, 8(2), 28414-28422.
- Losiana, L., Sari, S., & Yanto, Y. (2023). *Attitude of Media Tribunbengkulu.Com in Reporting on Fuel Price Policy*. *ISO Journal: Journal of Social, Political and Humanities Sciences*, 3(1), 105-118. <https://doi.org/10.53697/iso.v3i1.1255>
- Mantalean, V. (2023, September 14). Kompas.com Top of the Most Read Media Survey in the IPWS Version, Winning in 8 Categories.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.
- Michael Barlow. (2004). *MonoConc Pro: Concordance Software*.
- Noer, K. U., Kartika, T., Nurtjahyo, L. I., & Damaiyanti, V. P. (2022). Unraveling Sexual Violence in Higher Education: Preliminary Thoughts.
- Nurpadila. (2022). Women's Representation on the FTV Show Suara Hati Istri "Episode Does a Wife Receive Always Hurt" on Indosiar. <https://repository.iainpare.ac.id/>.
- Puspitasari, A., Wiyanti, E., & Nutriputra, I. (2022). Analysis of Critical Discourse on Social News of Kompas Daily Newspaper and Its Implications for Indonesian Language Learning. *Allegory: Journal of Indonesian Language Education Students*, 1(2), 248-258.
- Sugiyono. (2015). *Mixed Research Methods*. Bandung: Alfabeta.

- Supiandi. (2025). Media and Gender Stereotypes: A Feminism Study of Women's Representation in the Indonesian Entertainment Industry. *Ma'arif: Journal of Islamic Religious Studies and Philosophy*, 7(1), 1-17.
- Tarumingkeng, R. (2025). *Critical Discourse Analysis (CDA): Definition, Theory, and Application*. RUDYCT e-PRESS, 8.
- Tenriawali, A. Y. (2018). Representation of Victims of Violence in the Online News Text of the East Tribune: A Critical Discourse Analysis. *Totobuang*, 6(1), 1-5.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. John Benjamins Publishing.
- UN Women, 2022, *Gender Equality: Women's Rights in Review 25 Years after Beijing*.
- Wendratama, E. (2017). *ONLINE JOURNALISM*. Sleman, Yogyakarta: PT Bentang Pustaka.

Address for correspondence:
Shierly Amalia Aura Hayati
Faculty of Language and Arts
State University of Surabaya
amaliashierly@gmail.com

(Received 13 April 2025; revision received 20 November 2025;
Accepted 04 December 2025; final revision received 21 December 2025)