

---

# Understanding gender violence depiction in online news media

---

**Caitlin Loftus**

cloftus@uchicago.edu  
The University of Chicago

**Roberto Barroso-Luque**

barrosoluquer@uchicago.edu  
The University of Chicago

**Rukhshan Mian**

rukshshan@uchicago.edu  
The University of Chicago

## Abstract

Our study blabla

# 1 Introduction and Background

The declaration on the Elimination of Violence against Women, which was adopted by the United Nations in 1993, defines gender-based violence (GBV) as “any act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life”.

“Violence against women and girls is one of the most systematic and widespread human rights violations” according to the United Nations Entity for Gender Equality and the Empowerment of Women (UN Women). According to research by the World Health Organisation, in 2018 1 in 3 women worldwide had been subjected at least once in their lifetime to physical and/or sexual violence from any current or former husband or male intimate partner, or to sexual violence from someone who was not a current or former partner.

Public interest journalism has great potential to help in the fight against gender-based violence. A recent example of this was the MeToo movement, which largely began as a result of investigations by the New York Times. Despite the potential, the media’s portrayal of gender based violence can also have negative impacts on how society views these types of violence. A literature review found number of key themes in the way news media portray violence against women including “perpetuating myths and misrepresentations”, and “directly and indirectly shifting blame from male perpetrators of violence and assigning responsibility for violence to women”.

The media’s reporting can impact the public’s perception of GBV in a number of different ways. Firstly what is chosen to be reported on influences the types of violence the public are aware of and see as important issues. Secondly, the particular language choices made in the reporting—such as the words chosen to describe GBV, the structure of the writing, and particular linguistic devices—has a big impact on the public’s perception of GBV. For example, previous research has shown that the use of the passive voice in reporting of rape cases removes blame from the attacker (Bohner, 2001). Braber (2014) also demonstrated the use of the the passive voice in newspaper reporting about domestic violence cases, for example “an article in The Guardian describes the killing and the murder of a woman and her daughter, without mentioning the man who killed them.”

Why is this important/interesting to study using NLP techniques? MISSING MISSING MISSING

The aim of our work MISSING MISSING MISSING

For the purpose of our research, we look at Mexico, the UK and Pakistan. In recent years there have been a number of important cases and movements surrounding gender-based violence in these three countries. We include Mexico to understand how GBV movements have been reported in the news and see how political ideologies may impact reporting methods. Instances of GBV are prevalent in the UK as well. Sarah Everard’s death in early 2021 set off protests and movements against such instances. The case was reported on in various ways by different news outlets – some resorted to victim-blaming whereas some focused on more in-depth analysis. These patterns have been present in the past as well in terms of what kind of language is being used to depict instances of gender-based violence. Lastly, we take into account Pakistan where GBV is a widely underreported topic. Such instances are highly prevalent and a significant percentage of women experience domestic violence. Causes of underreporting include: A lack of belief in a legal system, lack of accountability in the (policing) authorities to actually report these issues, etc. Victim-blaming is unfortunately the norm in most instances and our goal is to analyze the sort of language news outlets use to describe such cases.

The motivation behind looking at passive voice instances in newspaper articles related to GBV stems primarily from pre-existing literature. It suggests how passive voice predominates in mass media reports describing male violence against women. This tends to put the actor in the background and the acted-upon person in the focus of discourse. An experiment from Germany found how passive voice positively correlated positively with rape-myth acceptance and perceived responsibility of the victim. This study was aimed at identifying subtle linguistic indicators of blaming victims of sexual violence, and at relating these to direct judgments of responsibility. Although the experiment was conducted on non-professional writers, we aim to extend this to look at how relatively professional writers describe cases of gender-based violence. We aim to build on existing literature and delve more into the extent to which passive voice is used in such newspaper articles.

Motivation for this topic also stems from anecdotal evidence. We often see cases of GBV being reported in newspapers. The language used in such articles often places the blame on women. Or rather, the language used intentionally tries to take away from the GBV act. An example of this can be seen here from an article published in the Sun, a right-wing newspaper from the UK: The language used is passive and the words being used to describe the instance itself are vague. Attention is being given to what the victim was doing as opposed to focusing on what the perpetrator did. Existing literature combined with such evidence lay the foundation for what we try to explore. Our aim is to look at the extent to which passive voice occurs in GBV-related newspaper articles in Mexico, UK and Pakistan.

First proposed by Harris, in the seminal work *Distributional Structure* (1954) and popularized by Firth (1957) the Distributional Hypothesis of Language suggests that words that occur in similar contexts have similar meanings and that differences in linguistic form connote differences in meaning. This linguistic theory has given rise to a plethora of subsequent research by which Natural Language Processing methods have been used to understand the semantic meaning of words and concepts in different corpora. Two recent studies are of particular relevance to our research.

Kozlowski et al. (2019) made use of the exciting method of word embeddings to analyze millions of books published during the span of 100 years and track how the meaning and social context of class shifted amid social transformations. The authors trained word embedding models on Google Ngrams text from books published over the span of the twentieth century to gain insights into socio-cultural understandings of social class. Surprisingly, this work found that word embeddings which include stereotypes and other harmful biases tend to accurately represent the cultural systems and corresponding text which give rise to such stereotypes. In addition, the authors find that the meaning of terms as represented by word embeddings closely resemble the meaning of these same terms in human responses, based on responses in the cloud-source platform Mechanical Turk.

Caliskan et al. (2017) use Global Vector representations, or GloVe embeddings to understand historical biases in text across a multitude of dimensions including race and gender. This work finds that there exists a relationship between prejudicial human behavior and the division between ingroup-outgroups which can be understood through the semantic relationships between words in language.

Taken together, these two papers exploit the assumptions of the Distribution Hypothesis of Language to gain insights into complex socio-cultural concepts by creating multidimensional vector representations of human language. We follow a similar methodology by using the Word2Vec algorithm and creating word embeddings of each of our compiled corpus, one language model for each newspaper. In doing so we seek to understand which terms are related to each other by comparing their cosine distance in multidimensional vector space. This analysis allowed us to understand and compare multiple cultural and social concepts and their relationship to domestic violence in contemporary online media.

## 2 Methodology

### 2.1 Web Scrapping

	Newspaper	Political ideology	Number of articles
Mexico	El Herald	Right	2820
	El Universal	Centre	3513
	La Jornada	Left	2099
Pakistan	Nation	Centre-Right	968
	Dawn	Centre-Left	330
	The News	Centre-Left	930
UK	The Times	Centre-right	1206
	The Sun	Right	1269
	The Guardian	Left	4568
Total			17778

Table 1: Article and Newspaper Counts

### 2.2 Word2Vec and Sentiment Analysis

We used the Word2Vec algorithm introduced by Mikolov et al (2013) to create word embedding representations of each of the newspaper corpus we compiled. The raw text from each document (e.g., an article for each newspaper in our dataset) was tokenized and normalized using word-tokenize and sentence-tokenize methods in Python. Sentence tokenized documents were retained in order to create a continuous bag of words (CBOW) representation of each newspaper corpus using Gensim’s implementation of the Word2Vec algorithm.

Once created, word embedding models for each of the newspapers were visualized using Principal component analysis and TSNE. PCA was used to create linear mappings of the multi-dimensional WE models by reducing each WE representation into the ten principal component vectors. The first two principal components were then plotted as well as a scree plot to understand the variance explained by each PC vector. Due to the linear limitations of PCA, we decided to employ TSNE to create a non linear low-dimensional representation of our high-dimensional WE spaces (Figures 1-6). Keywords were then selected and highlighted in both PCA and TSNE plots to understand their relationship to other keywords and the corpus in general. By observing the relative distance between keywords in these lower dimensional spaces we then made inferences of semantic meaning for each keyword and comparisons of their use across the political spectrum.

Furthermore, we made use of pre-trained neural network sentiment models in order to understand the relative polarity of keywords in each corpus. In order to get the average polarity of words in a specific corpus, we iterated over each keyword and over each WE model finding the closest words, based on cosine similarity, for each keyword  $i$  in each WE model  $j$ . We then used VaderSentiment’s pre-trained model to evaluate polarity for the ten “closest” words and assigned the average polarity to their respective keywords. Moreover, we used the same methodology with pre-trained Glove Embeddings (English) and the Spanish Billion Word Embeddings (Spanish) as points of comparison for the relative polarity of our keywords in their respective languages.

Finally, semantic algebra using the vector representation of specific words was used to create an analogy solver as proposed by Mikolov et al (2013). Using this methodology we investigated the analogy of woman: victim as man: ?? as described in (equation 1).

$$analogy(m : w \rightarrow k : ?) = \operatorname{argmax}_{v,w,k} (\cos(v, k) - \cos(v, m) + \cos(v, w)) \quad (1)$$

### 3 Results

#### 3.1 Semantic Meaning in Word Embedding Models

As described in our methodology section, we visualized the relationship between words in our corpus using PCA and TSNE. Specifically, we focused on the relationship of a set of keywords (equation 2) and their direct translation to spanish.

$$w \in \sqcup \text{man, woman, abuse, place, victim, blame, responsibility} \sqcup \quad (2)$$

Figure 1 and figure 2 show examples of PCA, TSNE and scree plots for a word embedding models fit to The Guardian (UK) and The Sun (UK) newspapers. We find that the relative distance between specific keywords such as woman, abuse and man in The Guardian WE model (Figure 1) when compared to the relative distance to the same three words in The Sun WE model (Figure 2) and The Times WE model (Supplementary Figure 5) show differential semantic use of these three words by each newspaper. The Guardian, which is left of center in the political spectrum, associates the word man as a middle word between abuse and woman as seen in the central position found in its WE model. In contrast, The Sun, the most conservative UK newspaper in our dataset, creates a much closer association between woman and abuse, placing man farther away as seen in the PCA plot for this newspaper. The Times, a moderate UK newspaper, follows a similar pattern to the Guardian. The finding in the UK using relative position of man, abuse and woman is not replicated neither in Pakistan nor in Mexico’s WE models. The appendix shows supplementary figures 4, 5 and 6 with 2-dimensional visualizations for all newspapers grouped by political tilt for all three countries. For both Pakistani and Mexican newspapers it is difficult to draw conclusions from the PCA and TSNE plots of their respective embeddings. Yet by using sentiment analysis and semantic algebra interesting patterns start to arise.

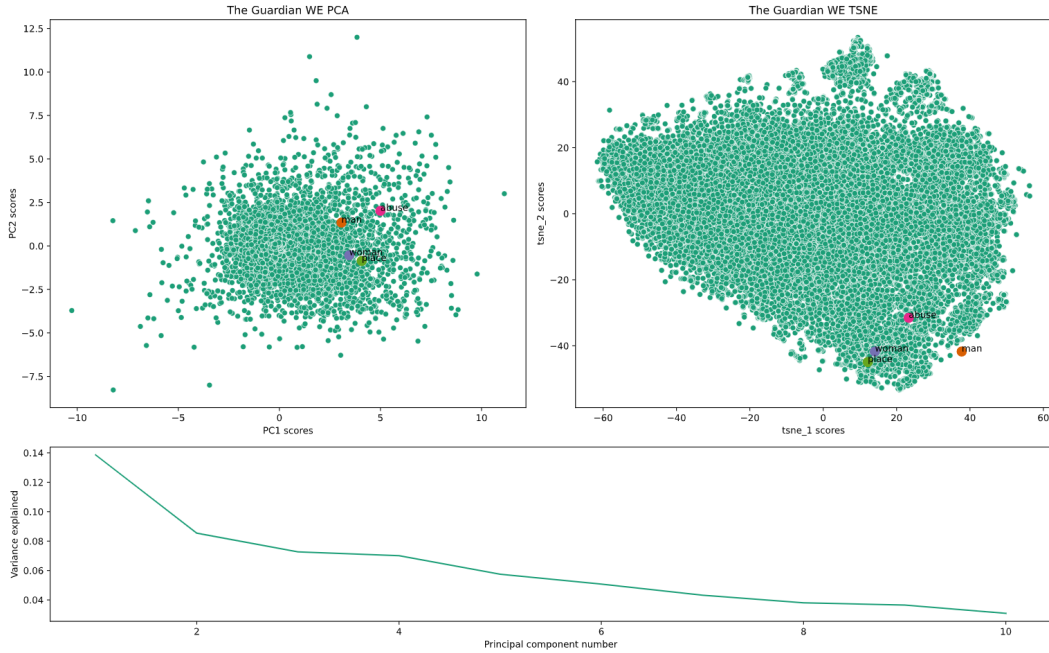


Figure 1: Visualization of embedding space (The Guardian UK)

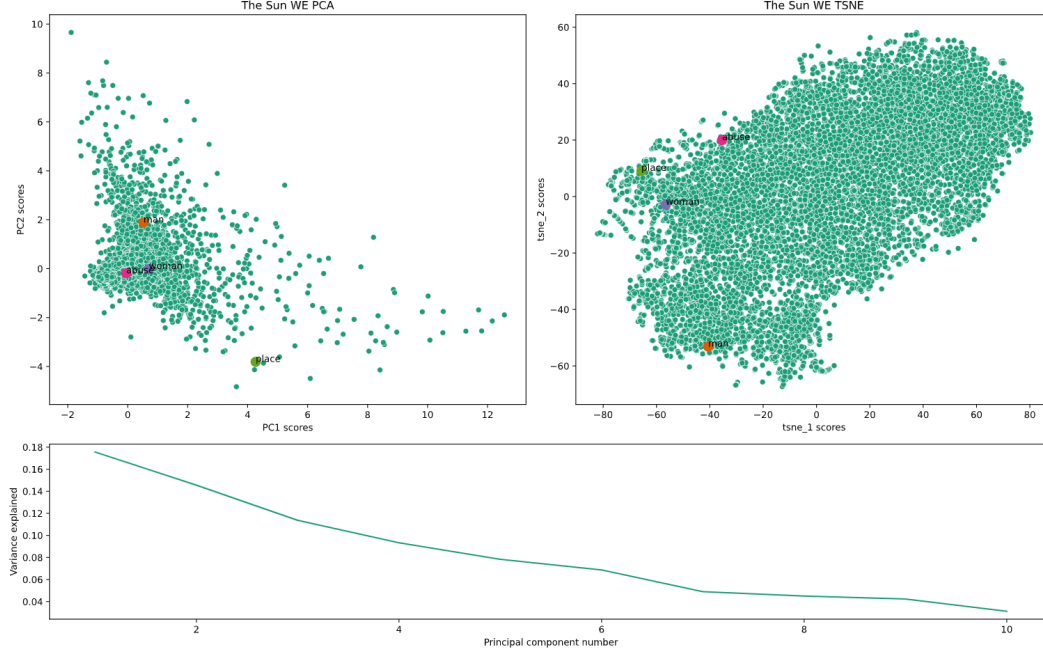


Figure 2: Visualization of embedding space (The Sun)

When using semantic algebra to create a word analogy solver (equation 1) we find a clear relationship between ideological tilt and semantic meaning for man and woman in media depictions of GBV. In table 2 we see the results of the analogy operation for the vector representations of woman, victim and man as described in (equation 3).

$$w_{woman} + w_{victim} - w_{man} = ?? \quad (3)$$

When focusing on the results for newspapers in Mexico, we find that the words that best solve the analogy for right wing newspaper el Herald are “expresion”, “condicion”, “completa” (complete). The solution to the same analogy for the moderate El Universal are “libre” (free), “crueldad” (cruelty), “erradicar” (eradicate), etc. While the solutions in La Jornada’s WE representation are “subordinacion” (subordination), “violenta” (violence), “castigar” (punishment), etc. In a similar fashion, when looking at the same analogy for UK newspapers, the words that best solve the analogy in The Guardian embeddings are words such as “perpetrator”, “abuser”, etc. while the words that best solve the analogy in The Sun embeddings are “domestic”, “harassment” etc. This pattern, which we consider one of the critical findings in our work, suggests that when solving for analogies, moving from right to left in the political spectrum, there is an increased association of a blame/attacker dimension attributed to man in the lexical space for both Mexican and UK newspapers. This finding is not replicated in the WE representation of Pakistani newspapers.

	Newspaper	Political ideology	Result
Mexico	El Herald	Right	(‘expresión’, 0.91), (‘condición’, 0.90), (‘completa’, 0.89), (‘ideología’, 0.88), (‘conllewa’, 0.88)
	El Universal	Centre	(‘libre’, 0.92), (‘política’, 0.92), (‘manpowergroup’, 0.92), (‘crueldad’, 0.92), (‘erradicar’, 0.92)
	La Jornada	Left	(‘definición’, 0.82), (‘subordinación’, 0.82), (‘violenta’, 0.81), (‘mujerla’, 0.81), (‘castiga’, 0.80)
Pakistan	Nation	Centre-Right	[(‘domestic’, 0.99), (‘harassment’, 0.99), (‘country’, 0.99), (‘increase’, 0.99), (‘rise’, 0.99)]
	Dawn	Centre-Left	
	The News	Centre-Left	
UK	The Sun	Right	(‘survivor’, 0.75), (‘domestic’, 0.75), (‘harassment’, 0.73), (‘demand’, 0.71), (‘mutilation’, 0.71)
	The Times	Centre-right	(‘charge’, 0.99), (‘offence’, 0.99), (‘allege’, 0.99), (‘violence’, 0.99), (‘report’, 0.99)
	The Guardian	Left	(‘survivor’, 0.71), (‘perpetrator’, 0.53), (‘domestic’, 0.49), (‘abuser’, 0.48), (‘stalker’, 0.44)

Table 2: Semantic Algebra:  $w_{woman} + w_{victim} - w_{man} = ??$

Finally, we go on to use pre-trained sentiment scoring models to find the average polarity of words in our keyword list for each newspaper in our dataset. In addition, we plot the results for the same analysis using the ubiquitous Glove embeddings and Spanish Billion Word (SBW) embeddings. Figure 3 highlights the results of this analysis. One interesting finding is the polarity relationship between newspaper embeddings and the topic agnostic Glove and SBW embeddings. In both Mexico and Pakistan, we find high variance across all newspapers and no obvious correlation between the relative polarity of any one newspaper with Glove or SBW. Nonetheless, in the UK we find that, on average, the polarity of keywords in The Guardian’s WE model tends to mirror, for most words, the polarity found in Glove embeddings. These results might be due to noise artifacts in our data generating process or could provide insights into the general semantic style employed by each newspaper when reporting on GBV. Because of the limitations in inferring semantic meaning in the respective depiction of GBV in our dataset we sought to further understand our corpus through passive/active voice detection.

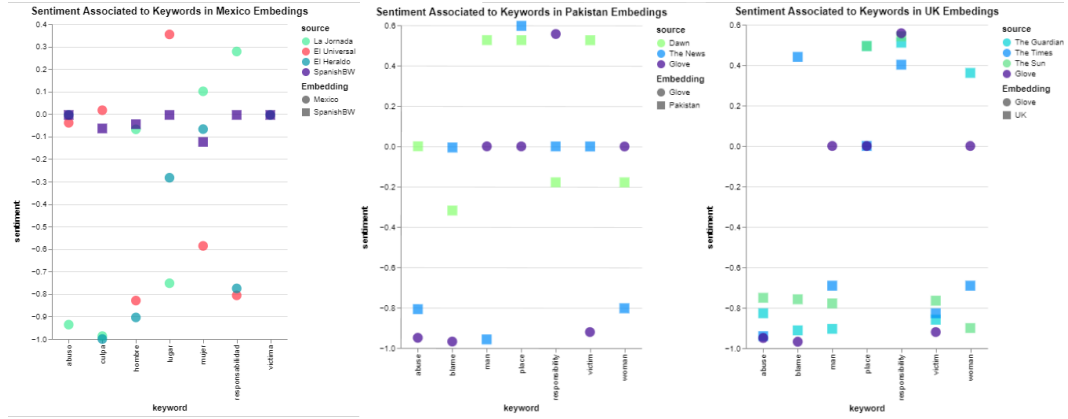


Figure 3: Sentiment towards keywords

### 3.2 Active Passive Voice

MISSING MISSING MISSING

### 3.3 LSTM Binary Classifier

	Test set accuracy	Test set F1 score
UK best model	0.78	0.83
Pakistan best model	0.83	0.88
Mexico best model	0.74	0.71

Table 3: Evaluation Metrics LSTM GBV Classifier

## **4 Conclusion, further research and limitations**

BLABLABLA





## 5 Appendix

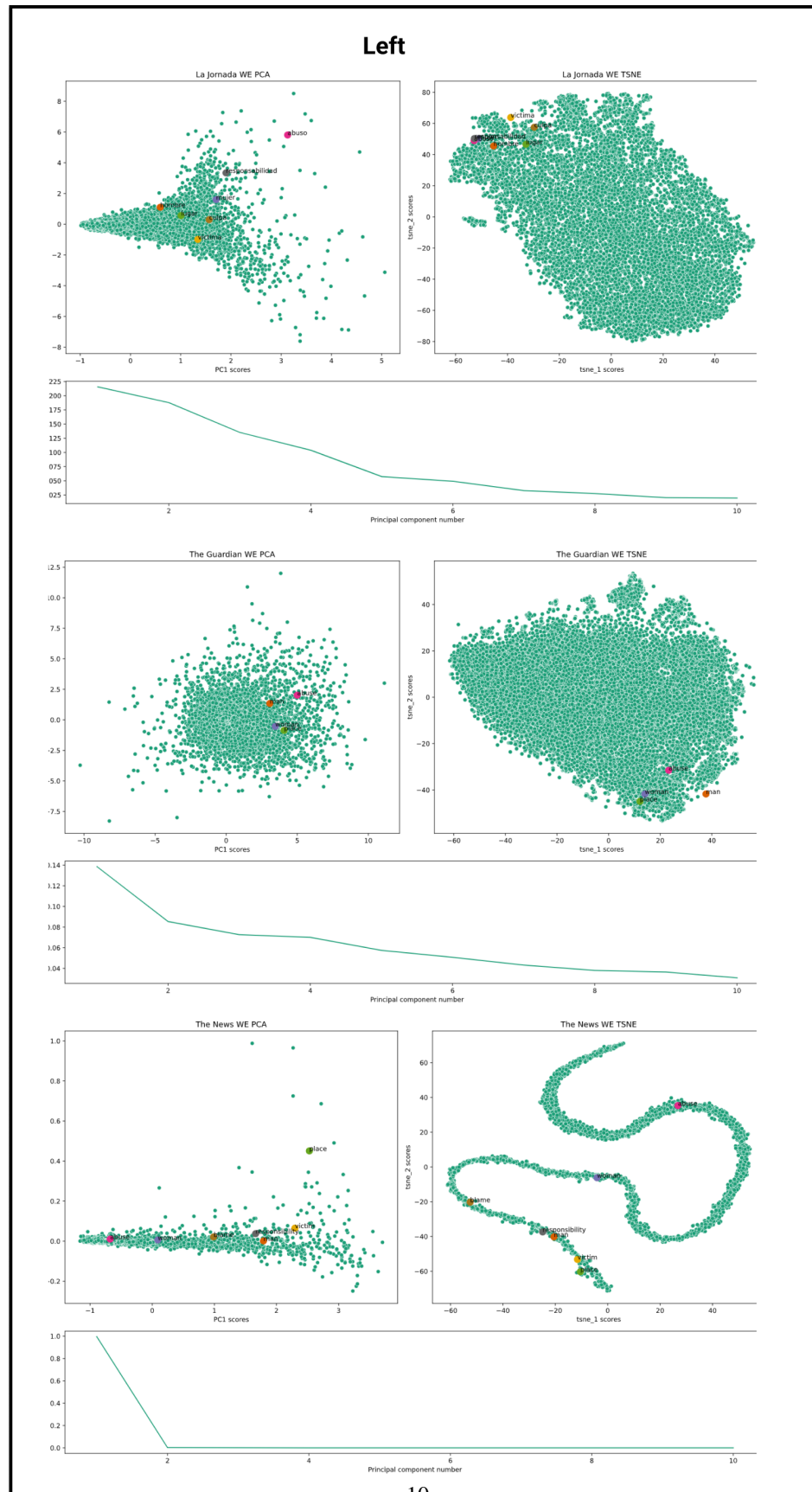


Figure 4: GBV depiction in left leaning media



### Conservative (right)

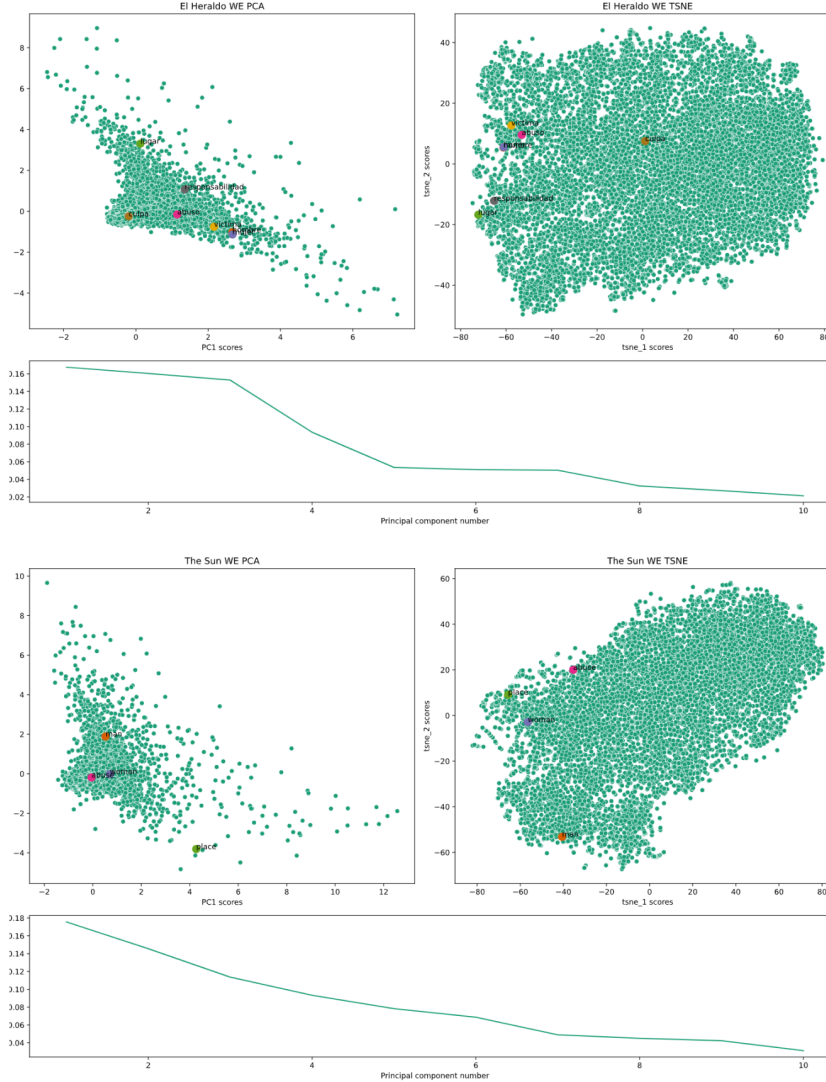


Figure 6: GBV depiction right leaning media

## 6 References

- [1] Bello, H. J., Palomar, N., Gallego, E., Navascués, L. J., and Lozano, C. (2020). Machine Learning to study the impact of gender-based violence in the news media. arXiv preprint arXiv:2012.07490.
- [2] Busso, L., Combei, C. R., and Tordini, O. (2020). Narrating Gender Violence A Corpus-Based Study on the Representation of Gender-Based Violence in Italian Media. In G. Giusti, and G. Iannàccaro (Eds.), *Language, Gender and Hate Speech : A Multidisciplinary Approach* (Language, Gender and Hate Speech A Multidisciplinary Approach). <https://doi.org/10.30687/978-88-6969-478-3/002>
- [3] , A., Bryson, J. J., and Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183-186. doi:10.1126/science.aal4230
- [4] Bail, C. A. (2012). The fringe effect. *American Sociological Review*, 77(6), 855-879. doi:10.1177/0003122412465743
- [5] Kozlowski, A. C., Taddy, M., and Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, 84(5), 905-949. doi:10.1177/0003122419877135
- [6] Sendén, M. G., Sikström, S., and Lindholm, T. (2015). “She” and “He” in news media messages: Pronoun use reflects gender biases in semantic contexts. *Sex Roles*, 72(1-2), 40-49. doi:10.1007/s11199-014-0437-x
- [7] Von Nordheim, G., Muller, H., and Scheppe, M. (2019). Young, free and biased: A comparison of mainstream and right-wing media coverage of the 2015–16 refugee crisis in German newspapers. *Journal of Alternative and Community Media*, 4(1), 38-56. doi:10.1386/joacm000421
- [8] Grimmer J. & Stewart B.M (2013): Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*
- [9] V. Rodelo, Frida, and Carlos Muñiz. “La Orientación Política Del Periódico y Su Influencia En La Presencia De Encuadres y Asuntos Dentro De Las Noticias.” *Estudios Sobre El Mensaje Periodístico*, vol. 23, no. 1, 1970, pp. 241–256., doi:10.5209/esmp.55594.
- [10] Harris, Zellig S. “Distributional Structure.” *Papers on Syntax*, 1981, pp. 3–22., doi:10.1007/978-94-009-8467-71.
- [11] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS’13)*. Curran Associates Inc., Red Hook, NY, USA, 3111–3119.