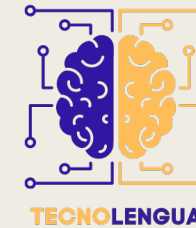


A LARGE NEWS CORPUS FOR THE DISCOURSE ANALYSIS OF VIOLENCE AGAINST WOMEN (VAW)

Miguel Fuster-Márquez [IULMA-UV]

Carla Fernández-Melendres [UMA]



IULMA

Institut Interuniversitari de
Llengües Modernes Aplicades
de la Comunitat Valenciana

Instituto Interuniversitario de
Lenguas Modernas Aplicadas
de la Comunidad Valenciana

Index

1. The NEWSGEN project
2. Principles & theoretical framework
3. Analytical methods
4. Objectives
5. Corpus design, compilation & annotation
6. NEWSGEN_VAW: Overview
7. Applications
8. Conclusions, limitations, and future perspectives

1. The NEWSGEN project

“News values and ideology: The discursive, cross-cultural construction of gender and social inequalities in (digital) press through online, real time corpus linguistic tools. *The Gender Gap Tracker and Kaleidographic.*”

Grant PID2019-110863GB-I00, funded by MCIN/AEI/10.13039/501100011033



Study and analysis of public discourses on **gender and social inequality** in the (digital) press

- ✓ Diachronic analysis
- ✓ Political, cultural, social and ideological impact



CADS

Frameworks & Methods:

- ✓ Discursive News Values Analysis
- ✓ Critical Discourse Studies
- ✓ Corpus-Assisted Discourse Studies

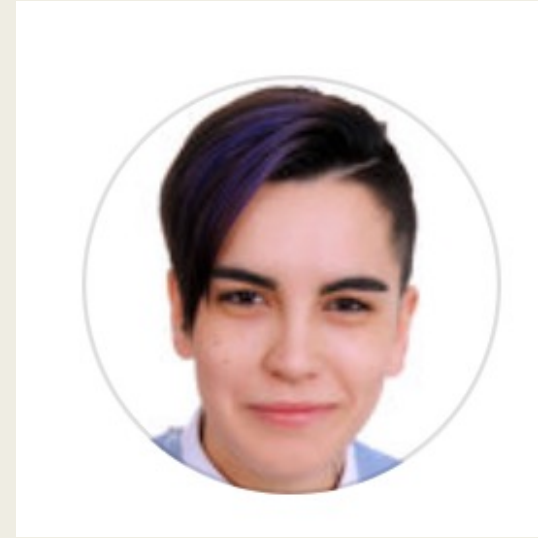


Corpora:

NEWSGEN_VAW (Violence Against Women)
NEWSGEN_DES (social inequality)

1. The NEWSGEN project: *TEAM*





1. THE NEWSGEN PROJECT: TECHNOLOGICAL ASSISTANCE (*TECNOLENGUA-UMA*)

2. Principles & theoretical framework: *Critical Discourse Analysis*

- An important assumption is that the press plays a crucial social role in shaping public opinion in, say, two opposing directions:
 - (1) reproducing (common/ shared norms, values, beliefs) or
 - (2) resisting discourses (i.e. challenging these social norms, values, beliefs) (Van Dijk 1991).
- The main focus is the analysis of *ideologies* and *inequalities* in news discourses, which already has a established tradition in CD studies.

2. Principles & Framework: views of news discourse & VAW

Constructivist view of discourse(s): calls into question the belief in objectivity or neutrality of news stories (cf. Burr 1995: 48)

CD analysts claim that all kinds of discourse have a purpose/goal, regardless of whether the linguistic (and any non-linguistic) choices are conscious or unconscious. The ideological expression in discourses is unavoidable.

Our interest is not related to the *linguistic analysis of news stories around* in itself, but in unveiling the *underlying ideologies linguistically construed* by journalists in such stories around VAW.

Of particular relevance to our VAW approach is found in works at the intersection between discourse and gender studies: e.g., in Butler (1990), Lazar (2005) or Ehrlich Meyerhoff and Holmes (2014), Maruenda-Bataller (2021), Santaemilia (2021), Fuster-Márquez (forthcoming)

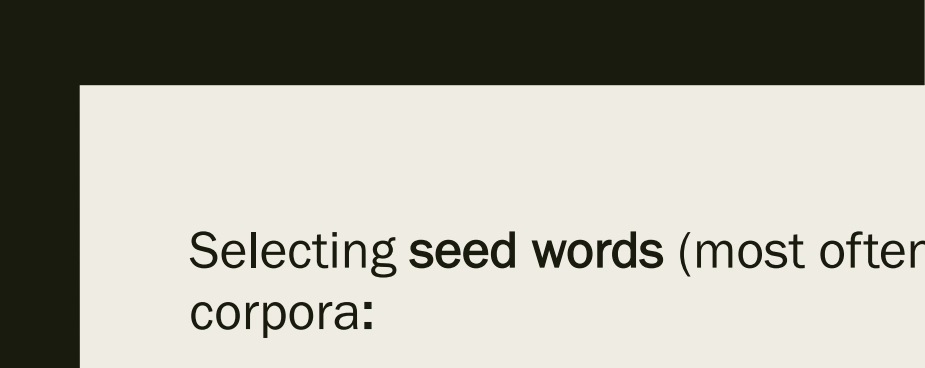
3. Methodology: CADS

- The compilation of a relatively large computerised corpus of this kind also implies that analysts borrow and take advantage of methods typically used in Corpus linguistics.
- The analysis of discourse by means of corpus linguistics techniques is known in the literature as **Corpus-assisted Discourse Studies**, described by Partington, Duguid and Taylor (2013: 10) as:
"that set of studies into the form and/or function of language as communicative discourse, which incorporate the use of computerised corpora in their analyses".


4. Objectives

Design, compilation, and annotation of a journalistic corpus (news stories and editorials) from the UK, the USA, Spain, and Catalonia around the topic of Violence Against Women (VAW).

Diachronic multilingual digital news corpus of Violence Against Women
(VAW)



Selecting **seed words** (most often the case in (Critical) Discourse Studies for ad-hoc corpora:

1. which cover the topic: Violence Against Women,
 2. In cross-linguistic and cross-regional studies, the seed words should be semantically and pragmatically equivalent in
 1. *three languages: English, Spanish, Catalan*
 2. *three countries: Spain, UK & US*
 3. *And should allow research for diachronic analysis*
- 

5. Corpus design, compilation & annotation: issues in corpus design



General questions concerning corpus compilation: **size, representativeness and balance** (Xiao 2010: 148: 153)

- **Size.** Bowker and Pearson (2002: 49) consider that there is no pre-established ideal number of words, since it depends on the purpose of study. Sinclair (1991: 18) believes that “a corpus should be as large as possible and should keep growing.”
- **Representativeness.** Defined by Biber (1993: 243) as “the extent to which a sample includes the full range of variability in a population”
- **Balance.** Sinclair (2005) “the proportions of different kinds of text in a corpus should correspond with informed and intuitive judgements.” Douglas (2003: 34) considers the balance of a corpus to be secondary to good practice.

5. Corpus design, compilation & annotation: issues in corpus design

Search TEXT: violencia de genero OR viol... DATE: All Dates SOURCE: El País - Nacional (Spain, S... MORE ▼

Dow Jones (0) All (7,466) **Publications** Web News (0) Blogs (0) Pictures (0)

Sort by: Relevance Duplicates: Similar

Date Export

Distribution: Yearly

Companies Export

Endesa, S.A.	55
Prisa - Promotora de Info...	50
United Nations	31
European Union	23
Fundación Anar	18
Spain Ministry of Health	17
Spain Ministry of Industry...	15
European Parliament	14
Siemens Gamesa Renew...	11
The Junta of Andalusia	11

Headlines 1 - 50 of 7,466 Next 50 Total duplicates: 31

- 1. Cuatro detenidos por la muerte de una pareja en Granada**

El País - Nacional, 23 April 2022, 295 words, Javier Arroyo, JAVIER ARROYO, (Spanish)

La Guardia Civil ha detenido a cuatro hombres por su implicación en la muerte de un hombre con signos de violencia en un barranco de la localidad costera de Sorvillán (524 habitantes) (Document PAISN00020220423ei4n0003j)
- 2. El 016 atiende a las víctimas de violencia machista las 24 horas y en 52...**

El País - Nacional, 20 April 2022, 64 words, (Spanish)

El 016 atiende a las víctimas de violencia machista las 24 horas y en 52 idiomas, al igual que WhatsApp en el 600 000 016. Los menores pueden dirigirse a la Fundación ANAR 900 20 (Document PAISN00020220420ei4k0002u)
- 3. Un hombre mata a su pareja y a su hija y se suicida en Girona**

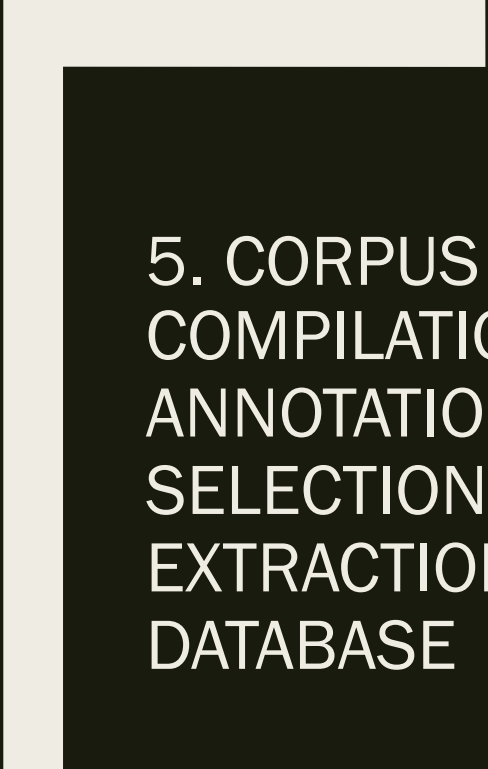
El País - Nacional, 20 April 2022, 465 words, Marta Rodríguez, MARTA RODRÍGUEZ, (Spanish)

Otro hijo de la fallecida, que vive en Francia, dio la voz de alarma Los Mossos d'Esquadra casa de la urbanización Els Pinars, en el municipio de Lloret de Mar (Girona). Las primeras (Document PAISN00020220420ei4k0002w)

+ 1 duplicate article(s) identified
- 4. La lucha de Mainer para "liberar" a su hija en Orán**

El País - Nacional, 19 April 2022, 950 words, M. ORMAZABAL, (Spanish)

Amira era un bebé de cuatro meses cuando en mayo de 2018 partió de España con sus padres vitoriana de 23 años, solo ha conseguido verla en los últimos cuatro años en contadas ocasiones (Document PAISN00020220419ei4j0002w)



5. CORPUS DESIGN, COMPILATION & ANNOTATION: SELECTION & EXTRACTION FROM DATABASE

5. Corpus design, compilation & annotation: appearance in FACTIVA

EL PAÍS

España

Cuatro detenidos por la muerte de una pareja en Granada

Javier Arroyo

JAVIER ARROYO

295 words

23 April 2022

El País - Nacional

PAISN

N (Nacional)

21

Spanish

(c) Copyright DIARIO EL PAIS, S.L. <http://www.elpais.es>

La Guardia Civil ha detenido a cuatro hombres por su implicación en la muerte de un hombre y una mujer cuyos cuerpos fueron hallados anteayer con signos de violencia en un barranco de la localidad costera de Sorvillán (524 habitantes, Granada), según el delegado del Gobierno en Andalucía, Pedro Fernández. La investigación sigue su curso y no se descartan nuevas detenciones, según la misma fuente. Los fallecidos son un hombre de 61 años y su pareja, una mujer de 47, que fue concejal del PP en la localidad vecina de Gualchos-Castell de Ferro entre 2015 y 2019.

La jueza Irene Navarrete Cánova, titular del juzgado de Instrucción 1 de Motril, decretó ayer el secreto de sumario. El barranco de Los Yesos, en el que se encontraron los cuerpos, está cerca del cortijo en el que ambos residían. Fuentes de la investigación habían descartado desde el inicio que el caso tenga relación con la **violencia machista**.

La tarde del jueves, día en el que fueron hallados los cadáveres, la línea de investigación prioritaria vinculaba las muertes con un ajuste de cuentas por drogas. El fallecido ha sido identificado como J. A. V., apodado el Rey León, a quien se relaciona con el tráfico y plantación de cáñamo, según informa Efe.

La investigación se inició a primeras horas de la mañana del jueves, cuando la Guardia Civil recibió la llamada de la hija del hombre fallecido informando de la desaparición de la pareja y del hallazgo del vehículo de su padre en la salida de Castell de Ferro de la autovía A-7, cercana al domicilio, con restos de sangre. Pocas horas después, los agentes encontraron los dos cadáveres, semidesnudos, en el barranco.

Diario AS. SL

Document PAISN00020220423ei4n0003j

WATCH ME

DESCUBRE MÁS



Selección: ESPAÑA

SUSCRIBETE

INICIAR SESIÓN



EL PAÍS

España

ANDALUCÍA · CATALUÑA · COMUNIDAD VALENCIANA · GALICIA · MADRID · PAÍS VASCO · ÚLTIMAS NOTICIAS

SUCESOS >

Cuatro detenidos por la muerte de una pareja hallada en un barranco de Granada

La mujer fallecida fue concejal del PP en la localidad de Gualchos-Castell de Ferro entre 2015 y 2019

JAVIER ARROYO

Granada - 22 ABR 2022 - 14:33 CEST

NEWS STORY AS FOUND ONLINE

JAVIER ARROYO

Granada - 22 ABR 2022 - 14:33 CEST



Miembros de la Guardia Civil investigan la muerte violenta de un hombre y su mujer en Sorvilán (Granada), este jueves. ALBA FEIXAS (EFE)

La Guardia Civil ha detenido a cuatro hombres por su implicación en la muerte de una pareja en Sorvilán (524 habitantes, Granada), según el delegado del Gobierno en Andalucía, Pedro Fernández. La investigación sigue su curso y no se descartan nuevas detenciones, según la misma fuente. Los fallecidos son un hombre de 61 años y su pareja, una mujer de 47 años que fue concejal del PP en la localidad vecina de Gualchos-Castell de Ferro entre 2015 y 2019. Sus cuerpos fueron hallados este jueves [en un barranco de Los Yesos](#), anejo de la localidad costera granadina de Sorvilán este jueves.

La jueza Irene Navarrete Cánova, titular del juzgado de Instrucción 1 de Motril, ha decretado este viernes el secreto de



NEWSLETTER



Recibe EL PAÍS de la mañana con el mejor análisis de la actualidad

NEWS
STORY AS
FOUND
ONLINE (2)

España

Cuatro detenidos por la muerte de una pareja en Granada

Javier Arroyo

JAVIER ARROYO

295 words

23 April 2022

El País - Nacional

PAISN

N (Nacional)

21

Spanish

(c) Copyright DIARIO EL PAIS, S.L. <http://www.elpais.es>

La Guardia Civil ha detenido a cuatro hombres por su implicación en la muerte de un hombre y una mujer cuyos cuerpos fueron hallados anteayer con signos de violencia en un barranco de la localidad costera de Sorvillán (524 habitantes, Granada), según el delegado del Gobierno en Andalucía, Pedro Fernández. La investigación sigue su curso y no se descartan nuevas detenciones, según la misma fuente. Los fallecidos son un hombre de 61 años y su pareja, una mujer de 47, que fue concejal del PP en la localidad vecina de Gualchos-Castell de Ferro entre 2015 y 2019.

La jueza Irene Navarrete Cánova, titular del juzgado de Instrucción 1 de Motril, decretó ayer el secreto de sumario. El barranco de Los Yesos, en el que se encontraron los cuerpos, está cerca del cortijo en el que ambos residían. Fuentes de la investigación habían descartado desde el inicio que el caso tenga relación con la **violencia machista**.

La tarde del jueves, día en el que fueron hallados los cadáveres, la línea de investigación prioritaria vinculaba las muertes con un ajuste de cuentas por drogas. El fallecido ha sido identificado como J. A. V., apodado el Rey León, a quien se relaciona con el tráfico y plantación de cáñamo, según informa Efe.

La investigación se inició a primeras horas de la mañana del jueves, cuando la Guardia Civil recibió la llamada de la hija del hombre fallecido informando de la desaparición de la pareja y del hallazgo del vehículo de su padre en la salida de Castell de Ferro de la autovía A-7, cercana al domicilio, con restos de sangre. Pocas horas después, los agentes encontraron los dos cadáveres, semidesnudos, en el barranco.

Diario AS. SL

Document PAISN00020220423ei4n0003j

5. CORPUS DESIGN, COMPILATION & ANNOTATION: CLEANING TEXTS AND REMOVAL OF DUPLICATES

- Year of publication: encoded as four digits (YYYY)
- Name of country or region
- Name of newspaper
- A numeric ID code to identify each article: each of the files has a unique identification code
- Example: 2020_SP_ABC_0145.txt
(Year_country/region_newspaper_ID code)

5. Corpus design, compilation & annotation: Classification, labelling, and post-processing of corpus files

5. Corpus design, compilation & annotation: annotation

- Leech (1997:2) defines corpus annotation as “the practice of adding **interpretative, linguistic** information to an electronic corpus of spoken and/or written language data”.
- In TXT & XML, ready to use in Sketch Engine

- **Metadata:**

- <country>
- <newspaper>
- <date>
- <author>
- <news_section>
- <editorial>
- <wordcount>
- <language>
- <headline>
- <body>

```
<xml>
<article country="UK" newspaper="The Guardian" date="2000-01-03" year="2000" month="January" day="03" author="Diane Taylor" news_section="No section" editorial="left-centre" wordcount="805 words" language="English">
  <section section_type="headline">
    Go it alone.
  </section>
  <section section_type="body">
    Ann Hudson finally found the courage to leave her abusive husband after 26 years of marriage [...]
  </section>
</article>
</xml>
```

6. NEWSGEN_VAW: overview

Country	Source	Years	Texts	Tokens	Words	%
USA	Boston Globe	1987-2020	9,065	9,421,687	7,883,912	11.9
	New York Times	1980-2020	9,243	11,544,100	9,659,914	14.6
	Washington Post	1977-2020	7,992	10,044,341	8,404,935	12.7
Corpus USA			26,300	31,010,128	25,948,761	39.2
UK	Daily Telegraph	2000-2020	2,209	1,686,959	1,431,972	2.1
	The Guardian	1986-2020	10,729	15,287,037	12,976,453	19.3
	The Times	1985-2020	4,817	4,485,931	3,807,884	5.7
Corpus UK			17,755	21,459,927	18,216,309	27.1
SP	ABC	1997-2020	13,056	7,288,477	6,312,549	9.2
	El Mundo	1995-2020	8,126	6,420,942	5,561,175	8.1
	El País	2001-2020	11,955	7,457,408	6,458,865	9.4
Corpus SP			33,137	21,166,827	18,332,589	26.8
CA	Diari Ara	2013-2020	2,580	2,051,241	1,756,476	2.6
	El Periódico	2014-2020	2,681	1,801,817	1,542,895	2.3
	La Vanguardia	2013-2020	2,584	1,553,528	1,330,284	2.0
Corpus CA			7,845	5,406,586	4,629,655	6.9
TOTAL			85,037	79,043,468	67,127,314	

Database



Study Corpus:
VAW (seeds)

Formats:
.txt and .xml

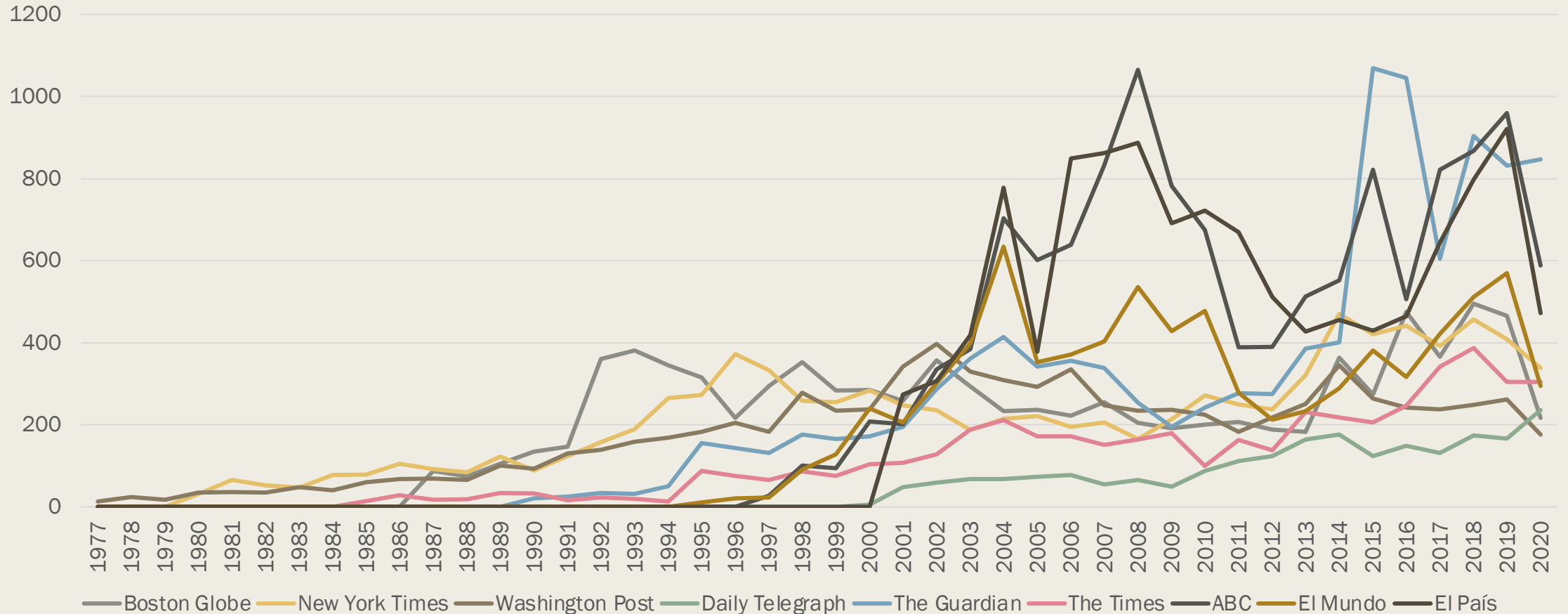
Metadata:

Country/region	Editorial
Newspaper	Wordcount
Date	Language
Author	Headline
Section	Body

6. NEWSGEN_VAW: overview (II)

USA
UK
Spain

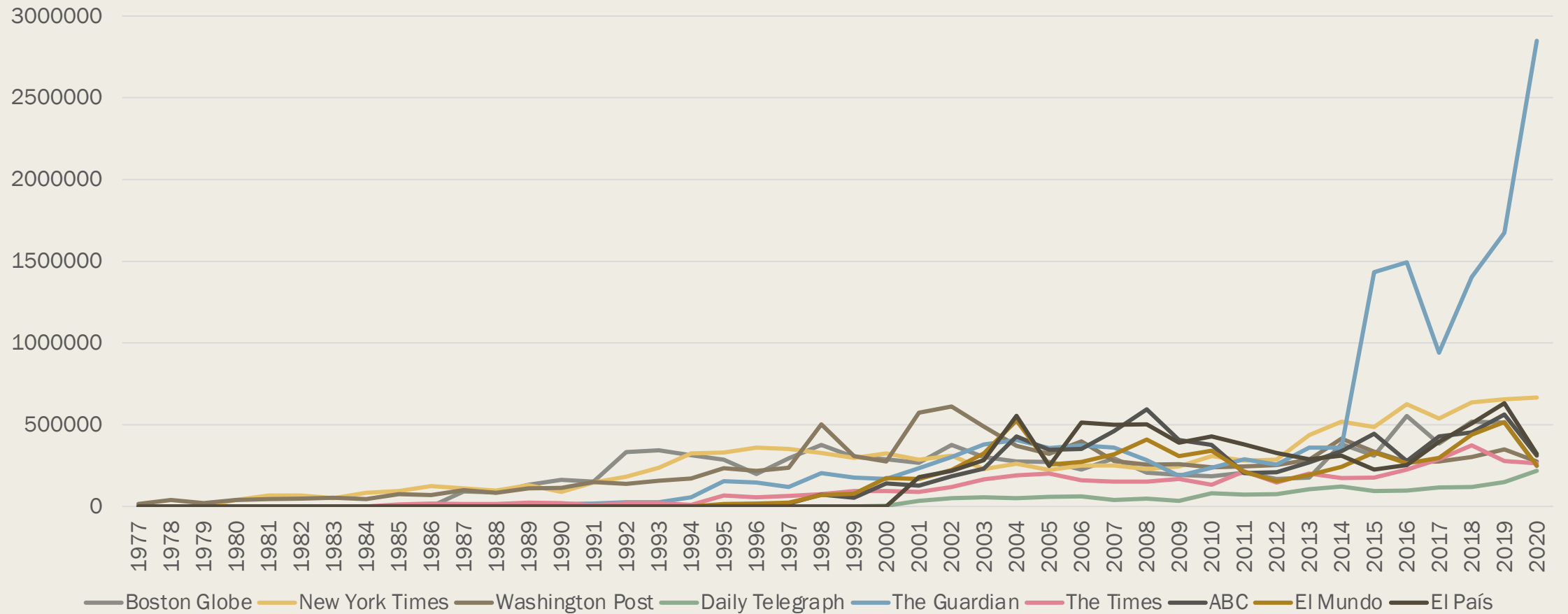
NEWSGEN_VAW
News articles



6. NEWSGEN_VAW: overview (III)

- USA
- UK
- Spain

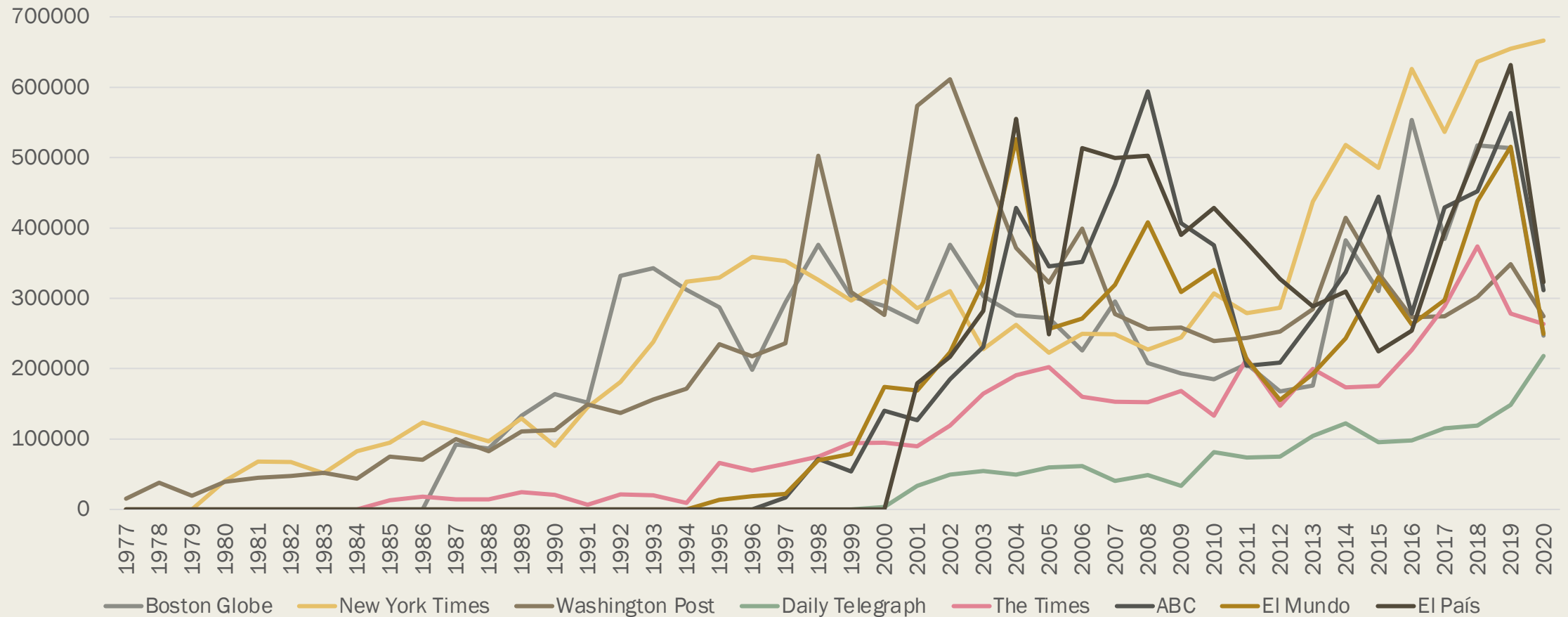
NEWSGEN_VAW
Tokens



6. NEWSGEN_VAW: overview (IV)

USA
UK
Spain

NEWSGEN_VAW
Tokens (without *The Guardian*)

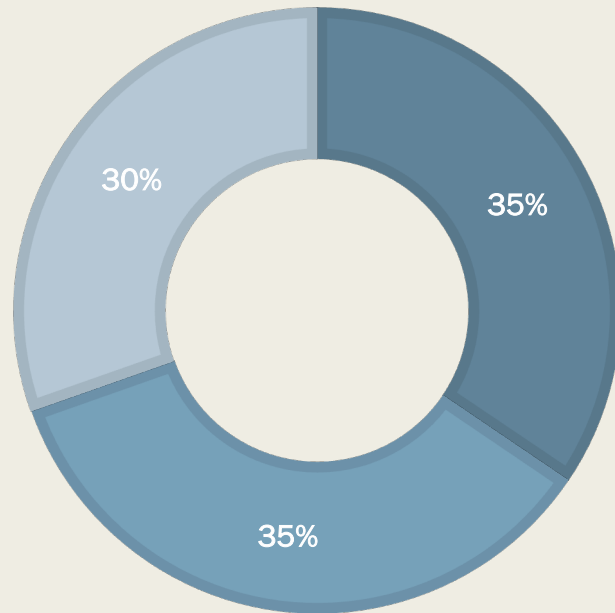


6. NEWSGEN_VAW: overview (V)

USA

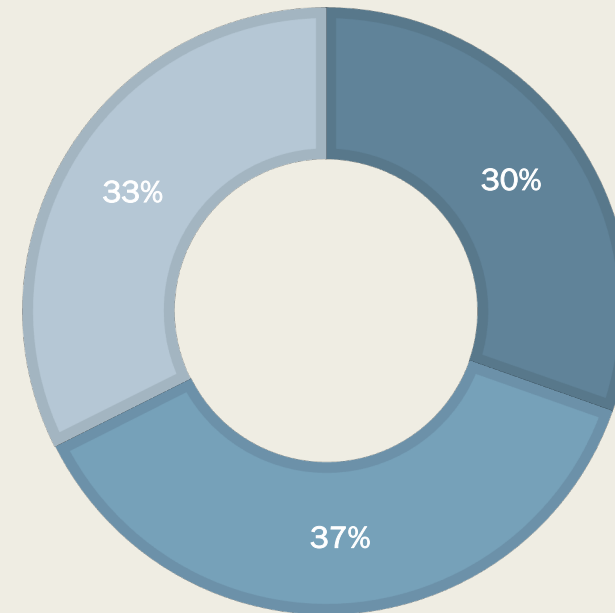
ARTICLES

■ Boston Globe ■ New York Times ■ Washington Post



TOKENS

■ Boston Globe ■ New York Times ■ Washington Post

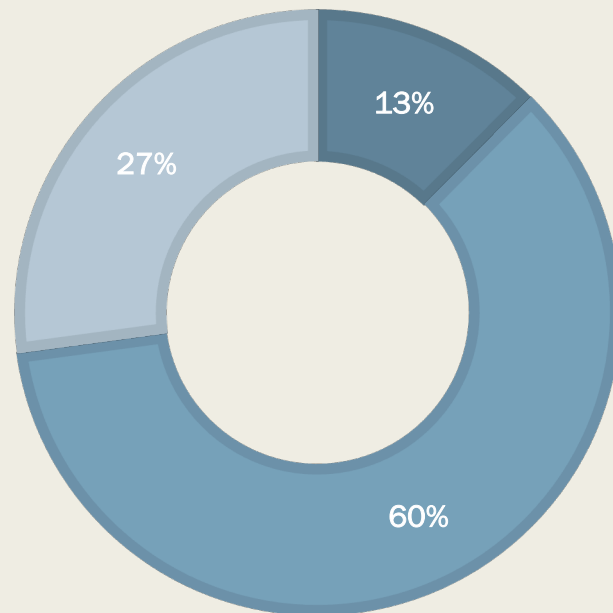


6. NEWSGEN_VAW: overview (VI)

UK

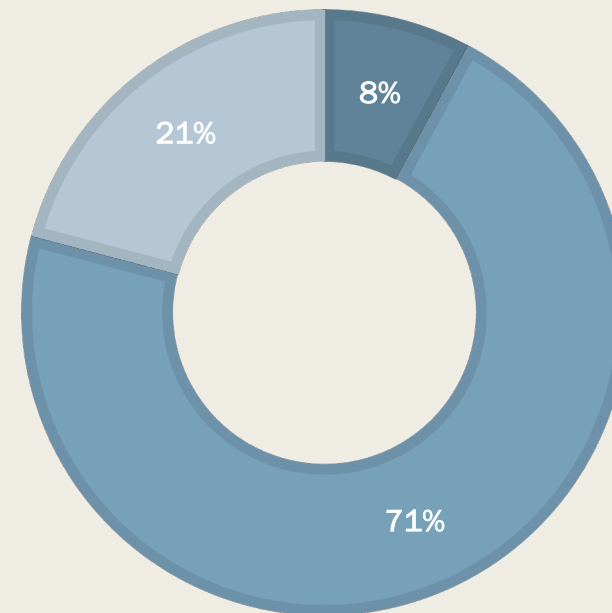
ARTICLES

■ Daily Telegraph ■ The Guardian ■ The Times



TOKENS

■ Daily Telegraph ■ The Guardian ■ The Times

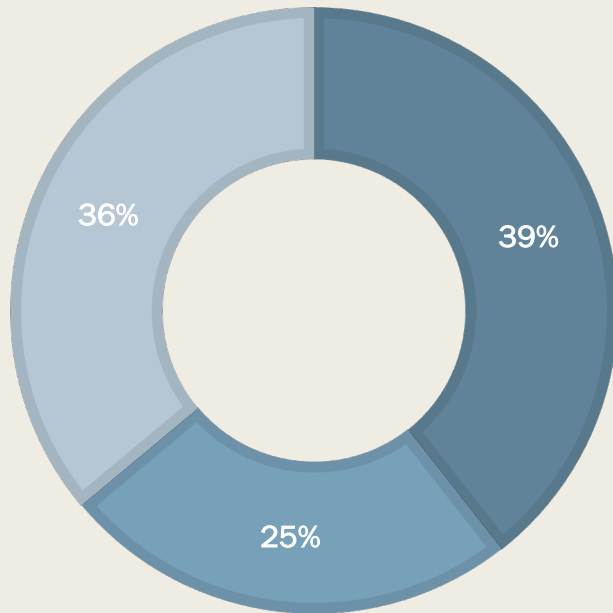


6. NEWSGEN_VAW: overview (VII)

Spain

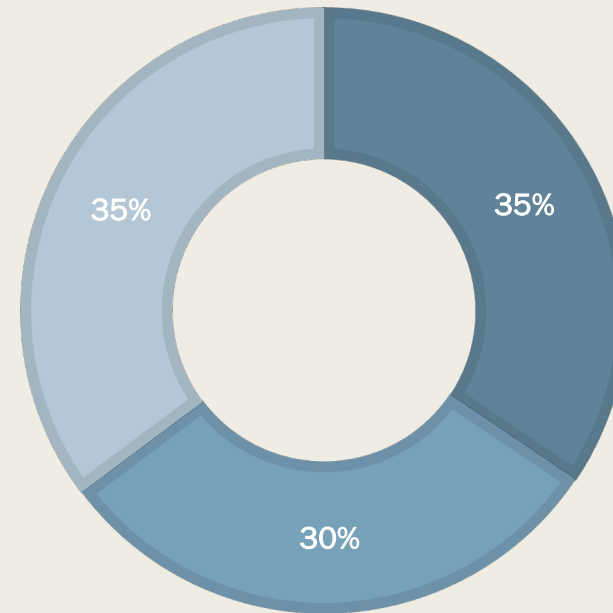
ARTICLES

■ ABC ■ El Mundo ■ El País



TOKENS

■ ABC ■ El Mundo ■ El País



NEWSGEN_VAW and NEWSGEN_DES are designed to examine aspects of **newsworthiness** but they allow for other kinds of **Corpus-assisted Discourse Analysis (quantitative and/or qualitative)**.

- Diachronic analysis
- Political, cultural, social, and ideological impact in news texts.
- Potential searches by selecting from the metadata: headlines, body, date, country/region, newspaper, author, section, editorial, length, language varieties (British English vs. American English) ...

7. Applications

7. Applications

Sketch Engine



- Sketch Engine provides us with tokenization, POS tagging, parsing and lemmatisation necessary to process our study corpus.
- Various techniques (concordances, word sketches, frequency lists, keyword lists, etc.) can be used together with in-build large corpora in Sketch Engine which allow for useful comparisons.
- Thanks to the annotation of our NEWSGEN corpora, and also word-level annotation in Sketch-Engine we can make use of these techniques at different levels of comparison.

7. Applications

Trends: change over time. *Daily Telegraph* (2000-2020)



Word	Trend ↓	Frequency	Sample	Word	Trend ↓	Frequency	Sample	Word	Trend ↓	Frequency	Sample
1 abusers	↗ 2.25	80		18 cost	↘ -1.28	137		35 support	↗ 0.93	646	
2 online	↗ 1.96	246		19 experiences	↗ 1.23	108		36 area	↘ -0.90	192	
3 opinion	↗ 1.96	82		20 killing	↘ -1.23	228		37 abuse	↗ 0.90	1,495	
4 era	↗ 1.88	69		21 round	↘ -1.19	134		38 either	↘ -0.90	197	
5 via	↗ 1.66	111		22 sport	↗ 1.19	177		39 red	↘ -0.90	125	
6 survivors	↗ 1.60	110		23 warned	↗ 1.19	273		40 girl	↘ -0.87	352	
7 battered	↘ -1.60	82		24 sexist	↗ 1.19	75		41 process	↗ 0.87	171	
8 campaigners	↗ 1.54	86		25 near	↘ -1.15	200		42 due	↗ 0.87	237	
9 murders	↘ -1.54	143		26 supporting	↗ 1.15	111		43 trial	↘ -0.87	398	
10 brilliant	↗ 1.48	77		27 decide	↘ -1.11	77		44 common	↘ -0.87	190	
11 killings	↘ -1.48	117		28 students	↗ 1.07	154		45 yesterday	↘ -0.84	753	
12 intended	↘ -1.43	89		29 passion	↘ -1.00	137		46 large	↘ -0.81	181	
13 murder	↘ -1.43	681		30 beaten	↘ -1.00	95		47 parents	↘ -0.78	527	
14 overseas	↗ 1.43	99		31 bought	↘ -1.00	82		48 change	↗ 0.75	399	
15 conflict	↗ 1.38	123		32 wives	↘ -1.00	120		49 include	↘ -0.75	246	
16 amid	↗ 1.33	106		33 charity	↗ 0.97	446		50 prepared	↘ -0.75	106	
17 economy	↗ 1.33	83		34 chief	↗ 0.97	462					

8. Conclusions: limitations, and future perspectives

Conclusions:

The research team is always present and aware of the procedures at every stage of the compilation process.

Every stage requires important decisions to be made so that the objectives of the team are met by the corpus that has been compiled.

The software which has to be used to explore the corpus is also a crucial issue.

8. Conclusions: limitations, and future perspectives

Limitations

- Morphosyntactic annotation is here provided by the software Sketch Engine (a paid licence is needed for its use)

Future perspectives of the research team

- The **NEWSGEN_VAW** corpus potential can still be enhanced (e.g. new features can be added as metadata, and more newspapers can be added, and updated)
- The **NEWSGEN_DES** is in the final compilation stage.
- Creating a **Gender Gap Tracker** (Maite Taboada) in the Spanish context, with static and more dynamic journalistic corpora to measure the bias of gender presence in news stories.

References

- Baker, P., Gabrielatos, C., Khosravini, M., Krzyzanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society*, 19(3), 273–306.
- Baker, P., T. McEnery & C. Gabrielatos (2013) Sketching Muslims: A Corpus Driven Analysis of Representations Around the Word 'Muslim' in the British Press 1998–2009. *Applied Linguistics* 34(3): 255-278.
- Bednarek, M. & H. Caple (2017) *The Discourse of News Values: How News organizations create newsworthiness*. OUP.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing* 8/4: 243–257.
- Bowker, L. & Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge.
- Butler, J. (1990) *Gender trouble: feminist and the subversion of Identity*. Routledge.
- Ehrlich, S., M. Meyerhoff & J. Holmes (2014) *The Handbook of Language, Gender & Sexuality*. Chichester: Wiley.
- Factiva [database]. Retrieved from www.factiva.com.
- Fuster-Márquez, M. & C. Gregori-Signes (2019) La construcción discursiva del turismo en la prensa española (verano de 2017). *Discurso & sociedad* Vol 13(2): 195-224.
- Fuster-Márquez, M., J. Santaemilia, C. Gregori-Signes & P. Rodríguez Abruñeiras (2021) Insights from corpus-assisted discourse analysis: Unveiling social attitudes and values. In Fuster-Márquez, M., J. Santaemilia, C. Gregori-Signes & P. Rodríguez Abruñeiras (eds.) *Exploring discourse and ideology through corpora*. In *Linguistic Insights*. Bern: Peter Lang, 7-16.
- Kilgarriff, A., V. Baisa, J. Busta, & M. Jakubicek (2014) The Sketch Engine: Ten Years On. *Lexicography* 1 (1): 7-36.
- Lazar, M. (2005) *Feminist Critical Discourse Analysis: Gender, Power and Ideology*. Palgrave-Macmillan.
- Leech, G. N. (1997). Introducing corpus annotation. In: Roger Garside, Geoffrey Leech and Tony McEnery (eds.), *Corpus Annotation*, (pp. 1–18). London/New York: Longman.
- Maruenda-Bataller, S. (2021) The role of news values in the discursive construction of the female victim in media outlets: A comparative study. In Fuster-Márquez, M., J. Santaemilia, C. Gregori-Signes & P. Rodríguez Abruñeiras (eds.) *Exploring discourse and ideology through corpora*. In *Linguistic Insights*. Bern: Peter Lang, 141-166.
- Partington, A., A. Duguid & C. Taylor (2013) *Patterns and Meanings in Discourse : Theory and Practice in Corpus-assisted Discourse Studies (CADS)*. John Benjamins Publishing.
- Potts, Amanda, Bednarek, Monika & Caple, Helen (2015) How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on Hurricane Katrina. *Discourse and Communication* 9 (2): 149-172.
- Santaemilia, José & Sergio Maruenda (2011) Building a comparable corpus (English-Spanish) of newspaper articles on gender and sexual (in)equality (GENTEXT-N): Present and future applications in the analysis of socio-ideological discourses. In María Luisa Carrió Pastor & Miguel Ángel Candel Mora (eds.) *Actas del III Congreso Internacional de Lingüística de Corpus - Las Tecnologías de la Información y las Comunicaciones: Presente y futuro en el análisis de corpus*. Valencia. Editorial Universitat Politècnica de València. 197-204
- Santaemilia, José & Sergio Maruenda (2013) “Naming practices and negotiation of meaning: A corpus-based analysis of Spanish and English newspaper discourse.” In Istvan Kecskes & Jesús Romero Trillo (eds.) *Research Trends in Intercultural Pragmatics*. Berlin: De Gruyter Mouton. 439-457.
- Santaemilia, José & Sergio Maruenda (2014) The linguistic representation of gender violence in (written) media discourse: The term ‘woman’ in Spanish contemporary newspapers *Journal of Language Aggression and Conflict* 2(2): 249-273.
- Santaemilia, José (2021) News values as evaluation. Main naming practices in Violence Against Women news stories in contemporary Spanish newspapers: *El País* vs. *El Mundo* (2005-2010). *RiCL* 9 (2): 90-113.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2005). Corpus and text – Basic principles. In Martin Wynne ed. *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxbow Books. <http://users.ox.ac.uk/~martinw/dlc/index.htm> (7 May, 2020.)
- Van Dijk, T. (2001). Critical discourse analysis. In D. Tannen, D. Schiffrin, & H. Hamilton (Eds.), *Handbook of discourse analysis* (pp. 352–371). Oxford: Blackwell.
- Xiao, R. (2010). Corpus creation. In Nitin Indurkha and Frederick J. Damerau eds. *Handbook of Natural Language Processing*. (pp. 147–165). Boca Raton: Chapman & Hall/CRC.

Thank you!

Miguel Fuster-Márquez [IULMA-UV]

Miguel.Fuster@uv.es

Carla Fernández-Melendres [UMA]

cfdz@uma.es

