62 years went through: An analysis of sexism in the lyrics of the most-listened-to songs in Spain

Laura Casanovas-Buliart [1,*] Priscila Alvarez-Cueva [2] and Carlos Castillo [1,3]

[1] *Dept. of Technologies of Information and Communication, Pompeu Fabra University, Spain*
[2] *Dept. of Communication, Pompeu Fabra University, Spain*
[3] *ICREA, Spain.*

(Dated: July 10, 2023)

Sexism against women remains an entrenched problem, manifested in contemporary cultural production worldwide. Since cultural production, and music in particular, can be understood as both a mirror for and a reflection of the society where it is inserted, the persistence of sexism in music might rather represent how sexist our society remains. Therefore, the present work aims to analyze the evolution of sexism towards women among the most-listened to music lyrics during the past six decades. Specifically, we look at the most listened to songs in Spain. Moreover, by following an intersectional perspective, the study considers whether the lyrics refer to racialized bodies or not to further problematize their relationship. To perform a large-scale analysis spanning the lyrics of more than two thousand songs released throughout a period of about six decades (1960-2022), we use automatic text classification based on manually labeled training data to categorize music lyrics between sexist/non-sexist, racialized/non-racialized, while detecting broad categories of sexism speech present in them. We note that the automatic classifier labels paragraphs of songs rather than the entire songs. The main findings show that sexism has always been present in song lyrics in Spain, and the presence of it has increased considerably in the means of listening to music in the last decade. Moreover, some markers of sexism, such as words that focus on pleasure, women, the body, and sensual movements, are found. Finally, it is shown that there is a relationship between racialized and sexist lyrics, problematized from an intersectional perspective. The outcomes of this research could have an impact on applications with the potential to alleviate the widespread prevalence of sexist biases in society, as well as to automatize both qualitative and quantitative analysis among cultural studies.

Keywords: music, sexism, lyrics, machine learning, natural language processing.

I. Introduction

Within contemporary cultural production, music has been argued as the most consumed product (Bennett, 2001) due to both the high spread of videoclips (Illescas, 2015) and, specially, considering its lyrics (Guarinos and Valdellós, 2020, Wright and Qureshi, 2015). Debates about the impact of music – in all its forms, have been going on and off for quite some time whereas its production can be considered as an archive full of social interpretations within the culture it is inserted to (Whiteley, 1992). Indeed, the different currents of postmodernity believe that "culture is reality" (Kotarba, 1994) and, in this sense, text can be deemed an ethnography (Van Maanen, 1995) – language being (in some cases) the author's behaviour itself (Bryant, 1982). Therefore, text might be a basic power that composes a particular notion of reality (Richardson, 1991).

Both postmodern and feminist views argue that music and words are elements that embody a state or condition (Swidler, 1996), that may either help to build a stronger self (e.g., Van Bohemen et al. (2018)) or, on the contrary, continue to perpetuate gender roles and stereotypes (e.g., Álvarez-Cueva et al. (2021)). Nevertheless, when considering text and music, violent metaphors should not be deemed simple figures of speech (Eisikovits and Buchbinder, 1997) since verbal acts can go beyond melodies up to discriminate and violate the rights and values of a person or group of people. In this sense, lyrics can create an everyday reality available for everyone, especially relevant when considering its potential to influence people's behavior (Walker, 1994), in particular due to its impact on emotions (e.g., Eze (2020)).

Studies have argued how listening to altruistic and caring lyrics may increase the interpersonal empathy, helping to improve behavior, well-being and pro-social thought (Alvarez-Cueva, 2022, Daykin, Daykin, Greitemeyer, 2009), while listening to aggressive lyrics songs can strengthen violent thoughts and hostile sentiments (Anderson et al., 2003), and perpetuate discrimination and sexism, specially towards women and LGBTQ+ groups (Dhoest et al., 2015, Vandenbosch et al., 2013). One way or the other, it is undoubtedly that music is present in everyone's lives, and so are the lyrics that accompany it. Hence, it is a matter of concern that messages behind lyrics are discriminatory. Indeed, to explore what lyrics tell about our society and reality still leads scholars to keep open the debate about its role among societies' identification processes.

* Correspondence concerning this article should be addressed to Laura Casanovas-Buliart. Email: laura.casanovas03@alumni.upf.edu

Moreover, when it comes to cultural and music studies, in particular music lyrics, some works have argued about the existence of gender and race bias from stereotypical portrayal of women, e.g., in kipsigis secular songs (Koskei et al., 2018), to the analysis of gender roles, gender stereotypes and women objectification in contemporary popular music (Rasmussen and Densley, 2017, Smiler et al., 2017a), as well as in African soundtracks (Oyeka, 2019) and in country music (Flynn et al., 2016). In fact, most of the racialized music genres such as rap, hip hop (Schneider, 2011, Smiler et al., 2017b) and lately reggaeton and trap (Ángeles Díaz, 2019, Cadena Ruiz and Ortiz Lema, 2018) are continuously under scrutiny. However, sexism and sexist portrayals – either through images or lyrics, are not limited to them.

Nevertheless, in the analysis of cultural production in general, and of music production in particular, regarding gender representations (e.g., sexism, sexualization, gender roles, etc.), most of the techniques used to perform these investigations involve content analysis and manual coding, further problematized by a critical perspective. In so doing, researchers generally include small samples to dig deeper argumentations towards those representations in music production. However, when dealing with large amounts of songs, those approaches are complex and challenging to be used (Domènech, 2018).

Goal. The main goal of the present study is to examine sexism over time in a large data collection. More specifically, this research explores sexism against women in music lyrics of the most listened to songs in Spain over the last six decades. Automatic labelling processes to detect sexism and racial references are developed from manually labeled data, and used to create a high-accuracy automatic process that is applied over more than two thousand songs. Our paper starts from a critical gender perspective that highlights power relations between sex and race among music lyrics.

Organization. The organization of this article includes a theoretical section where related work is outlined (§II), a description of materials and methods (§III), results (§IV), conclusions (§V) and limitations (§VI).

II. Theoretical Framework and Related Work

This section presents the theoretical framework on sexism that underpins our work (§II A) followed by previous work on large-scale analysis of music lyrics (§II B) and large-scale detection of gender bias in text (§II C). Altogether, this section sets the theoretical context and previous works that the present study builds on.

A. Sexualization and sexism among it: Rhythm and rimes

Within postmodernity (Hall, 2011), the sexualization of western culture theorization took place regarding cultural production (Attwood, 2009). Under its "umbrella," sexual portrayals - included those that allude to sexism, have been problematized. However, from a feminist perspective, sexism itself can be more complex than a "single brush" among cultural production, in particular when it comes to music studies, since sexualization practices may also help to transform the understanding of the self, such as the well-discussed self-sexualization linked with a more determinant individual (Gill, 2003, 2017). Therefore, a question remains: How to determine whether a portrayal belongs to one or the other? Indeed, due to the common overlapping of sexualization and sexism, some debates tend to obscure the differences that are included among both music videos and lyrics in this regard which, in turn, may impact differently and in a deeper level on both the self and the different processes of identification (Skeggs, 1997). Therefore, it is important to firstly set a line that helps to clearly differentiate sexualization than sexism.

Sexism can be considered as the preconception against women – and anything feminine related, with the belief that heterosexual men are superior in terms of ability, intelligence, among others, and thus that women – and by extension anything feminine, deserve less treatment or benefit, especially considering sex differences (U.N.O.D.C., 2018). Indeed, debates about music production, not only among scholars but in media (e.g., television, radio, and social networks) tend to be in line with the concern towards sexism – as a discrimination practice that historically implied violence, harmful acts, or stereotypes reinforcement (especially towards women), whereas sexualization - although it may also contribute to the previous ones, also deals with elements of determination, sexual representation, among others (Álvarez-Cueva et al., 2021).

> "Sexiness has become a consumer good. It is packaged to the female consumer through discourses of 'choice', 'autonomy' and 'liberation', creating a new female sexual subjectivity that celebrates female agency and empowerment through consumption" (Evans and Riley, 2014).

Nevertheless, sexism remains present nowadays at different levels, these are benevolent (Good, 2017), hostile (Hack, 2017) and ambivalent (Grubbs, 2017) as well as institutional (Capodilupo, 2017), interpersonal or internalized (Bearman and Amrhein, 2014). However, as in sexualization practices, sexism may no longer be enacted exclusively towards women, in the sense that when including an intersectional lens, other levels of sexism - and oppressions, may be highlighted.

Precisely, from a postfeminist perspective, the understanding of sexism has been discussed among cultural production, in particular regarding advertising and media studies. Therefore, not only its definition but significance have moved from different fields and criticisms over the years. Gill (2011) argued about how sexism operates through media while intersecting other axes of power:

> "[I]f we think about sexism not as a single, unchanging 'thing' (e.g., a set of relatively stable stereotypes), but instead reconceptualize it as an agile, dynamic, changing and diverse set of malleable representations and practices of power, how could it be anything less than urgent to have this term in our critical vocabulary?" (Gill (2011), p. 62)

This is, in fact, how the author sets the contemporary cultural scenario: "Sexism has not disappeared, but has taken on new forms" (Gill (2011), p.63). Indeed, culture changes over time (Hall, 1992) and as such, every change is interrelated with the transformations that society promotes from different fields and movements. As such, feminism did have an impact on cultural production and our understanding of sexualization, sexism and the self. But, the question remains on how sexism is still represented, assuming that – as said before, sexism has simply been reformulated in new ways.

Walter (2010) has discussed these new "guises" for old sexism resurgence among society as well. According to the author, society – and to some extent culture, focus the understanding of sexual allure on hypersexuality or, as Attwood (2009) argued, the sexualization of western culture. This said, we are not participating in the discussion towards a clear definition of self-determination or empowerment, but rather, to consider gender biases of sexism with an intersectional approach. The Wex Definitions Team (2020) refers to gender bias as an individual who experiences disparate treatment due to their genuine or apparent gender identity.Therefore, we can explore the construction and prevalence of gender bias and sexism in music lyrics, as well as its intersection with race: racialized sexism.

Intersecting sexism. Sexism and other ways of oppression work as a dynamic that compound and interact with each other, denoting that every experience is unique. Therefore, as argued by Crenshaw (1989), the various levels of difference may transform such experiences in ways that, intentionally or not, sustain levels of power and oppression convenient for the system. This process of exploring and problematizing levels of difference and its overlapping have been part of other studies, in particular, those that follow a gender perspective. "To call for a revitalization of a notion of sexism is not treat it as a stand-alone ideology, but rather to reassert its place in a thoroughly intersectional analysis and politics" (Gill (2011), p.67).

When exploring sexism among music, scholars have critically explored the relationship between its representation – either in portrayals or lyrics (e.g., Aubrey et al.

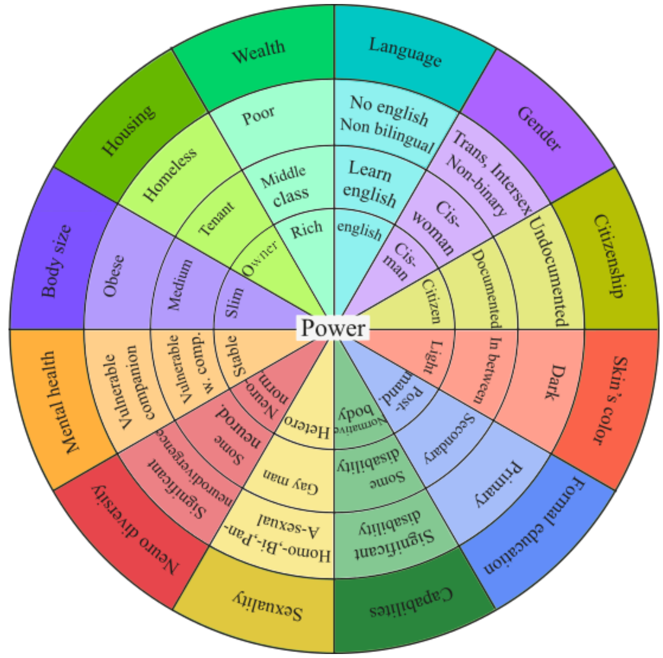


FIG. 1. Power wheel: The more centric a person characteristics are the more privileged this person is (own creation, adapted from `https://ccrweb.ca/en/anti-oppression`, inspired by `https://sylviaduckworth.com/`).

(2017), Eze (2020), Sollee (2015)). Whilst the focus of most studies are centred on women's representations, some also discussed how sexism can be directed toward men (Álvarez-Cueva et al., 2021), although racialization plays a key role when differentiating one portrayal from another. Indeed, the sum of oppressions interact among each other, meaning that both the disadvantages and privileges that a person has cannot be examined independently from other elements that compose the self.

The *power wheel* depicted in Figure 1 shows some of the multiple power relations that can be found in cultural production. As it represents, the further from the center, the more disadvantages the person will experience. For instance, a literate white man with a place to live and certain wealth, is more privileged than a black man with the same attributes; similarly, a white woman is more disadvantaged than a white man, but is more privileged than a black woman.

Nevertheless, as shown by other studies, most of the work towards sexism and music has been conducted by a qualitative approach or limited quantitative samples due to the complexity of mixed methods in this field. However, computational analysis methods have been formerly used with lyrics to complement traditional or manual analysis.

B. Large-Scale Analysis of Music Lyrics

When dealing with a large number of units of analysis, either music lyrics or video clips, manual analysis becomes economically unfeasible due to its high cost.

Mahedero et al. (2005) evaluated Natural Language Processing (NLP) methods applied on music lyrics to identify language, structure, themes or similarities between songs, and even genre classification by using Bidirectional Long Short-Term Memory (BLSTM) (Araújo Lima et al., 2020). Emotion recognition from music lyrics has been achieved using text processing methods such as word embeddings (Ara and Gopalakrishna, 2020), as well as detecting text reuse and similarities between artists (Meinecke and Jänicke, 2021). Explicit lyrics detection (lyrics not children appropriate) has been accomplished using FastText, BERT approaches and even logistic regression (Rospocher, 2021, 2022).

Barman et al. (2019) was the first to analyze bias and style using computational techniques on a huge dataset comprising five decades of music lyrics. They found that popular songs are quite different from other songs in terms of lyrics style, which indicated that lyrics is a key aspect when determining the popularity of a song. Furthermore, they observed that bias in lyrics correspond to bias in humans. To do this analysis they used a technique named WEAT with word2vec (Church, 2017) and FastText. Later, other studies have arisen following a similar structure and procedures to determine gender bias in songs (Gupta et al., 2021).

C. Large-Scale Gender Bias Analysis in Text

Regarding specific Natural Language Processing (NLP) tools, word embeddings are widely used to detect sexism and gender bias across different types of text, from legal documents to social media. There are studies (Grosz and Conde-Cespedes, 2020) that show that to detect sexism in workspace a good approach would be using GloVe (Pennington et al., 2014) and LSTM. However, for sexism detection in law documents the WEAT approach, which computes the association of word lists that represent possibly biased issues to a set of pronoun or otherwise gendered pairs (Caliskan et al., 2017), has been used and proved to outperform current NLP bias detection methods (Gillis, 2021).

Many studies have focused on detecting sexism in social media, which has increasingly become a concerning issue. Actually, the most common approach to detect sexism in text is using BERT approximations (Butt et al., 2021, Samghabadi et al., 2020). For instance, recently, a model (Frenda et al., 2019) for detecting misogyny in tweets has been described. As well as a neural architecture based on BERT, which carried out the first work on multi-labeling classification for sexism detection (Parikh et al., 2019). Moreover, there are in-depth analysis of sexist tweets that classified them as "hostile," "benevolent" or "others" (Jha and Mamidi, 2017), using Support Vector Machines (SVM), sequence-to-sequence models and FastText classifier (Joulin et al., 2016). (Sharifirad and Matwin, 2019) used different types of word embeddings mixed with LSTM and Naive Bayes models to detect harassment of different classes.

Language changes over time, as new words appear, other words die, and the existing ones adopt additional meanings. As it has been shown in previous studies (Garg et al., 2018, Kim et al., 2014, Kozlowski et al., 2019), changes in embeddings follow closely demographic shifts over time. Hence, word embeddings can be a powerful tool to quantify trends in social change, concretely they can be used to determine how gender stereotypes and points of view toward ethnic minorities evolve (Garg et al., 2018). Furthermore, there are studies (Charlesworth et al., 2021) that show that even in very different corpus of language gender stereotypes emerge consistently and robustly in all of them.

III. Materials and methods

Figure 2 overviews our data processing, which includes data collection (§III A), lyrics extraction (§III B), manual labeling and modeling (§III C), and automatic labeling (§III D).

A. Data Collection

The first step involved data collection, where research was conducted to identify the most popular songs from each decade from 1960 to 2022 in Spain, resulting in a dataset containing the song title, artist, hit year, language, and source. Radical changes in listening habits during the observation period meant that different data sources were needed for different years. In this study, we used three sources for retrieving the most listened-to songs in Spain: the lists of "Los40," "Los Superventas" and Spotify top songs. Currently, people in Spain listen to music an average of 20 hours per week.[1] Indeed, most people listen to music on streaming platforms such as Spotify, TikTok or YouTube, although there are people who listen to it on the radio or by other means. Moreover, the digital music market in Spain has increased immensely in the last years, being now much more important than the physical one. [2] Therefore, for the most recent period (2010-2022), the data was obtained using the Spotify API.[3] From 2010 to 2015, as well as for 2022,

[1] https://www.promusicae.es/descarga-informes/engaging-with-music-2021-infografia-datos-de-espana-n237/

[2] https://www.promusicae.es/descarga-informes/mercado-musica-grabada-espana-2019-n204/

[3] https://developer.spotify.com/documentation/web-api/

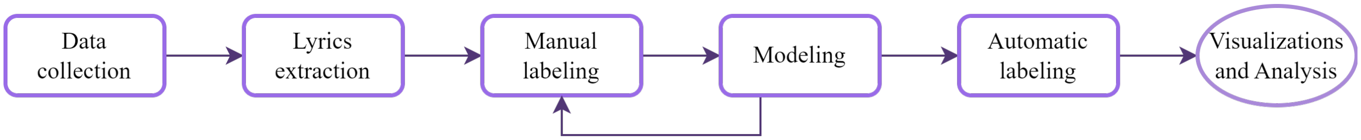


FIG. 2. In this figure, the data processing stages are shown: data collection, lyrics extraction, manual labeling, modeling, automatic labeling, and visualization and analysis. An extra loop between manual labeling and modeling is included since this process was done iteratively.

there were lists available of the top 100 tracks in Spain each year that could be downloaded.[4] Meanwhile, the playlists for the songs from 2017 to 2021 were available on a weekly basis.[5] To find the top songs of each of those years, we automatically identified the songs that were repeated with the most frequency throughout the year. This allowed us to obtain the most streamed songs of each year. During the decades before 2010, the radio was the main channel that people in Spain used to listen to music. The democratization of radio in Spain transformed it into a powerful tool for promoting popular music. Initially, Spanish music dominated during the thirties, forties, and fifties. However, "Discomanía" and the rise of rock and roll from the 1960s onwards reshaped radio, paving the way for it to become a major disseminator of this new genre. Charts like "Los Superventas" section (Pintado, 2017) directly influenced the musical preferences of the youth (Eguizábal Jiménez, 2021). "Los 40 Principales" (Los40) (LOS40, 2018), launched in 1966, initiated the music radio network in Spain. It pioneered the "Top Forty" style-like program and introduced innovative radio formulas, leading to the inauguration of exclusive stations in 1979. By 1988, Los40 surpassed other radio stations with nearly five million daily listeners, solidifying its position as the most popular radio station in Spain (Eguizábal Jiménez, 2021). Hence, the data for the oldest songs (from 1960 to 2009) was obtained from radio charts. Specifically, the data for the period from 1960 to 1965 was obtained from "Los Superventas," while the data for the most popular songs between 1966 and 2010 was sourced from "Los 40 Principales." The obtained sample consisted of a total of 2910 songs.

B. Lyrics Extraction

The second step was downloading the lyrics for the most popular songs identified. The majority were obtained using the Genius API (Genius Media Group, 2021), which is the largest collection of musical expertise and lyrical content worldwide, and the ones not found in the platform were downloaded manually. We completed the final analysis with 98.97% of the lyrics. The rest were either not located through manual Internet searches or were excluded during the automation process due to to unconventional formats. Hence, the final sample to be analyzed consists of 2,840 song lyrics.

Lyrics were parsed to divide them into paragraphs, which usually correspond to verses and choruses in popular songs. A total of 34,030 paragraphs were obtained. Our basic classification unit were these paragraphs.

[4] https://open.spotify.com/playlist/11H3xmArMqDAQf5sv7i8eD

[5] https://spotifycharts.com/regional/es/weekly/2016-12-23--2016-12-30

TABLE I. Pre-existing labeled datasets with sexist/not sexist categories, used for the preliminary classification model.

Dataset	Source	#Entries	Lang.
Lyrics	Slim (2019)	532	EN
Reddit comments	Guest (2021)	6567	EN
Tweets (Exist) TRAIN	EXIST (2021)	6977	EN, ES
Tweets (Exist) TEST	EXIST (2021)	4368	EN, ES
Tweets (MeTwo)	Rodriguez (2020)	3333	ES

C. Manual Labeling and Modeling

The process of manual labeling and modeling was undertaken in an iterative way, creating progressively more accurate models. To make labeling more efficient, each version of the model was used to stratify the sample of data to label for the next version.

Pre-existing labeled datasets. We acquired pre-existing datasets, summarized on Table I, with text labeled as sexist or not sexist. Most of these datasets are based on texts found in social media, such as Twitter postings ("tweets") or Reddit comments, which are quite different from music lyrics. These datasets were used exclusively to train a preliminary classification model for detecting sexism, not for the final model. We found only one dataset comprised of lyrics which we used for testing our preliminary classification model. Nevertheless, the lyrics in this dataset were not organized into verses or paragraphs.

New labels: sexist lyrics. A preliminary examination of the music lyrics allowed us to build a list of common ways in which sexism manifested: role and attribute stereotyping, body-shaming, hyper-sexualization, threats, sexual assault, sexual harassment, victim blaming, slut-shaming, motherhood-related discrimination, physical violence and gaslighting (Parikh et al., 2019). We also consider women objectification, which is seeing and/or treating a woman, as an object (Papadaki, 2010); rape, defined as a sexual aggression involving penetration (Art. 192 Spanish Penal Code); and control and possession behavior that is the sexist thoughts of men being in the right of controlling women or owning them. Control and possession behavior includes paternalism, which is the ideology that teaches men to minimize women's agency (Ware, 2020). Examples from these categories can be found in Table II.

In this research *"sexism"* refers to sexism against women, unless specified otherwise. Moreover, when detecting sexism in lyrics we also take into account the concept of romantic love, which is the misbelief that anything can be done and be justified for love (Nava-Reyes et al., 2018). Note that for this research, when we talk about racialized people, we refer to people whose ethnics have been mentioned when referring to that person (Schaefer, 2008).

Learning scheme. Our model is a supervised classification model that receives as input a paragraph of music lyrics, and outputs a binary inference on whether it contains or does not contain sexist language. While dynamic word embeddings such as BERT (Devlin and Chang, 2018) are sometimes more effective than other methods (Rudolph and Blei, 2018), for this model we chose to use Language-Agnostic Sentence Representations (LASER) embeddings (Artetxe and Schwenk, 2019). This is because LASER embeddings can be applied across multiple languages, and our sample is multilingual. As previously stated, we examined the performance of various models on several labeled pre-existing datasets. The Support-Vector Machine (SVM) (Cortes and Vapnik, 1995) classifier and the Long Short-Term Memory (LSTM) neural network classifier yielded the most favorable outcomes. We selected LSTM for its demonstrated effectiveness in text classification (Butt et al., 2021, Grosz and Conde-Cespedes, 2020, Sharifirad and Matwin, 2019), as evidenced in the section §II.

To simplify the task of text understanding for the models, we broke down the songs into smaller sections, referred to as paragraphs. These paragraphs were the ones that were consistently labeled, not the entire songs. Only at the end of this process, after determining the number of sexist or non-sexist paragraphs within each song, were the songs themselves labeled. A song was considered sexist if it contained at least one paragraph labeled as sexist.

The source code of the model can be found in the following anonymized repository: `https://gitfront.io/r/user-9355771/U8KmMaL3JSb3/sexism-in-lyrics/`

Iterative modeling and labeling process. We created three versions of the model. The first version was based on the pre-existing labeled datasets summarized on Table I. We used this version to sample paragraphs from songs from the years 1960-1969 (664 samples), and 2021 (1451 samples), which provide a diverse range of vocabulary. The sample was stratified by probability of being sexist, with roughly one third high probability, one third low probability, and one third medium probability. This hierarchization was instrumental in revealing the initial model's inability to accurately categorize paragraphs. In light of this, manual labeling was employed. This manual annotation involved labeling the paragraphs as sexist/non-sexist and racialized/non-racialized, and indicating the specific sexism category present. For that, the established criteria for determining whether a paragraph was sexist or not, which is outlined in section §I, was followed. This is, a paragraph was considered sexist if it fit into one of the analytical categories of sexism depicted on Table II against women. If there was any doubt as to whether a paragraph was sexist or not, it was not labeled as such in order to avoid confusion in the training process. Moreover, if a paragraph contained any reference to a person's ethnicity, it was classified as racialized and annotated accordingly; otherwise, it was deemed non-racialized. Annotations were done by one author of this paper and borderline cases discussed with another author of this paper.

The second version of the model was created using only these new labels, discarding the labels from previous work, which involved domains different from the one under study. We used it to sample, again in a stratified manner, about 100 randomly selected paragraphs from each decade (600 paragraphs in total). Furthermore, due to the limited number of positive samples of racialized references, a list of keywords was used to extract other potential samples and 826 paragraphs were manually labeled as either racialized or non-racialized. This information was then used to train the model that would detect racialization in lyrics. We added these samples to our training set, and used it to train the third version of the model.

Among the paragraphs labeled as sexist in the final version of our training set, we can find the categories of sexism shown in Figure 3.

D. Automatic Labeling

With the final dataset, we trained the final model, which has an AUC (Area Under the Curve) score (Hanley and McNeil, 1982) of 92% and F1 score (Van Rijsbergen, 1983) of 84%. Both number indicate a fairly high level of accuracy. We then used this model to annotate the remaining, unlabeled paragraphs.

TABLE II. Categories of sexism found in our music lyrics (own elaboration), from most frequent to least frequent in the manually-labeled training set

Category	Lyrics Examples
Control and possession behaviour	Baby I'm preying on you tonight.
Hypersexualization	Un movimiento muy sexy (sexy). *(A very sexy movement.)*
Attribute stereotyping	Tú sabiendo que e' una diabla, diabla. *(You knowing that she's a demon, demon.)*
Women objectification	Tu cuerpo a mí me aloca. Quiero por siempre tenerte, tenerte. *(Your body drives me crazy. I want forever to have you, to have you.)*
Body shaming	Con su trasero supo ganarse la admiración. *(With her ass she was able to earn admiration.)*
Physical violence	Solía amarla, pero tuve que matarla, tuve que enterrarla seis pies bajo tierra y aún puedo oír cómo se queja. *(I used to love her, but I had to kill her, I had to bury her six feet under and I can still hear her moaning.)*
Sexual assault, excluding rape	Que tengo la polla en candela y quiero comerte ese culo. *(I've got my dick on fire and I want to eat your ass.)*
Sexual harassment, excluding assault	Let me bend that ass over.
Slut shaming	Te veo enseñar el culo y las tetas a los tíos delante de mi cara. *(I see you flashing your ass and tits at guys in front of my face.)*
Victim blaming	Carolina trátame bien, o al final te tendré que comer. *(Carolina treat me well, or I'll have to eat you in the end.)*
Rape	Tendría que besarte, desnudarte, pegarte y luego violarte hasta que digas sí. *(I would have to kiss you, strip you, beat you and then rape you until you say yes.)*
Threats	Pues si quiero hacerte daño, solo falta que yo quiera. *(Well, if I want to hurt you, I only need to want to.)*
Gaslighting	Ella se hace la más boba. *(She plays the fool.)*
Role stereotyping	Ella sabe hacerlo todo en la casa. *(She knows how to do everything in the house.)*
Motherhood-related discrimination	Fuck you, hijueputa. *(Fuck you, motherfucker.)*

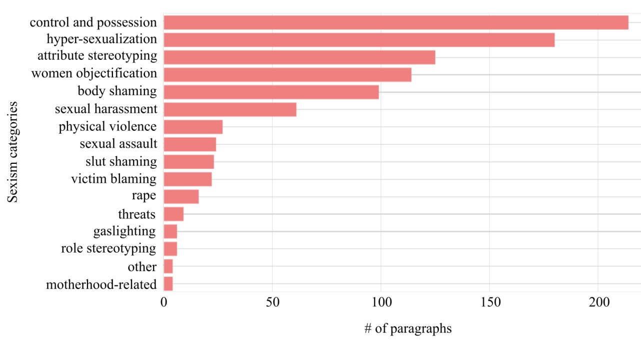


FIG. 3. Sexism categories in the manually-labeled training dataset

E. Crowdsourced data labeling

In parallel, we carried out two crowdsourcing tasks to gather information on how sexism is represented in lyrics. We focused solely on Spanish songs, as they were the most prevalent in sexist songs, in order to ensure that the future analysis would be more accurate and definitive. As aforementioned, the labeled dataset included a column with the categories of sexism present in each sexist paragraph.

The most prevalent categories in our labeled dataset were control and possession behavior and hyper-sexualization (Figure 3). As hyper-sexualization and women objectification have similar definitions by theory, we decided to combine them. To gain a deeper understanding of these frequent categories and determine if there were any characteristic words that would differentiate between them, we used crowdsourcing to collect more labels.[6] The participants were provided with a paragraph, a brief definition of the sexism category[7], and three examples from each category. The task was to classify the verses in the paragraphs as sexist or not based on the definition given.

The task was completed by three different people and the results were used to create two datasets, one for each category, containing verses that were labeled as belonging to the category. These datasets were then used to create word shifts visualizations (Gallagher et al., 2021), which are shown and analyzed in §IV D, to understand the words or phrases that mark these two sexism categories. It is important to note that in all cases, there was

[6] The platform used was Surge HQ, more information can be found in this link: `https://www.surgehq.ai/`

[7] Note that the defintion for the combination of hyper-sexualization and women objectification was rephrased as follows: *Unwarranted focus on women physical aspects or sexual acts, or seeing and/or treating a woman, as an object or animal*

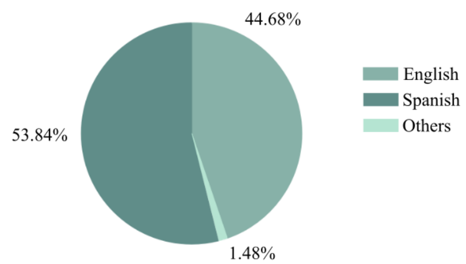


FIG. 4. Percentage of languages of all songs represented in the final dataset.

a minimum of two-thirds agreement between the participants on the final answer used in the study.

IV. Results

We start this chapter by presenting the prevalence of languages in our dataset (§IV A). Next, we show the percentages of sexist and racialized songs we obtained (§IV B). We expose some correlations and we study the markers behind sexism and racialization in lyrics (§IV D and §IV E). Later, we examine the prevalence of sexism in lyrics over time (§IV C).

A. Prevalence of language

For the purpose of the present study, we gathered song lyrics to categorize them into sexist or non-sexist. As shown in Figure 4, 54% of the most listened to songs in Spain from 1960 to 2022 were in Spanish, while 45% were in English, and the remaining 1.5% were in other languages (i.e., Catalan, French, German, Galician, Hebrew, Italian, Korean, Portuguese, Xhosa and songs with both Spanish and English). Hence, we will only be focusing on songs in English and Spanish separately in the following sections, as they constitute the majority of the sample. It is worth noting, however, that many Spanish songs include English words due to the hybridization of language in music within the context of post-modernity, as the music industry is primarily concentrated in the global north, particularly in the US and the UK.[8]

[8] 2021 music industry statistics: `https://gmr2021.ifpi.org/assets/GMR2021_State%20of%20the%20Industry.pdf`

B. Prevalence of sexism and racialization

In this section, we present the outcomes of the labeling process conducted on the lyrics of the most popular songs in Spain from 1960 to 2022, as illustrated in Figure 5.

Despite the possible common perception that only a negligible proportion of lyrics in the most listened songs in Spain over the past decades are sexist, our approach indicates that 51% of them are classified as sexist. In addition, out of the sexist songs, 60% are in Spanish and 38% are in English. Furthermore, the majority of songs that have between two and all of their paragraphs labeled as sexist belong to the 2010s to 2020s period, which reinforces the findings presented in section §IV C (e.g., Table III in Appendix VI). In addition, our analysis reveals that 92% of the featured artists in these songs are men. We also note that roughly 15% of these songs contain racialized paragraphs. Among the most prevalent genres in the songs from the last decades, *Urban Latin* stands out, although *Pop* is the only genre that appears across all decades. Conversely, we did a brief examination of 41 randomly selected non-sexist songs (see Table IV in Appendix VI for details), and we found that although Pop music remains dominant, other music genres represented are more diverse than in the sample of sexist songs. Anecdotally, we observe in the sample of sexist songs from Table III a larger prevalence of male artists than in the sample of non-sexist songs of Table IV.

In contrast, our results show that only 4% of the lyrics in the analyzed songs are classified as racialized, while 96% are considered non-racialized. Additionally, 93% of the racialized songs are in Spanish, with the remaining 7% in English. This finding suggests that racialization is more prevalent in the Spanish language than in the English language in Spain (refer to Figure 5). When comparing the categories of sexist/non-sexist and racialized/non-racialized songs, we found that around 8% of the sexist songs also fall under the category of racialized, while only 1% of the non-sexist songs are considered racialized. Since the number of racialized lyrics is limited, our ability to draw definitive conclusions is constrained. However, a significant proportion of racialized songs, around 89%, are also categorized as sexist, suggesting a potential relationship between racialization and sexism.

C. Sexism over time

As illustrated in Figure 6 a), the prevalence of sexism has risen in recent years. Aside from the peak in the 70s, the level of sexism remained relatively stable until the 2000s. However, the most concerning aspect of this study is the sharp increase in sexism over the last two decades, which has surged to 77%. This implies that 77 out of the 100 most frequently played lyrics are deemed sexist. However, as the last decade studied only comprises the years 2020, 2021, and 2022, we decided to examine the

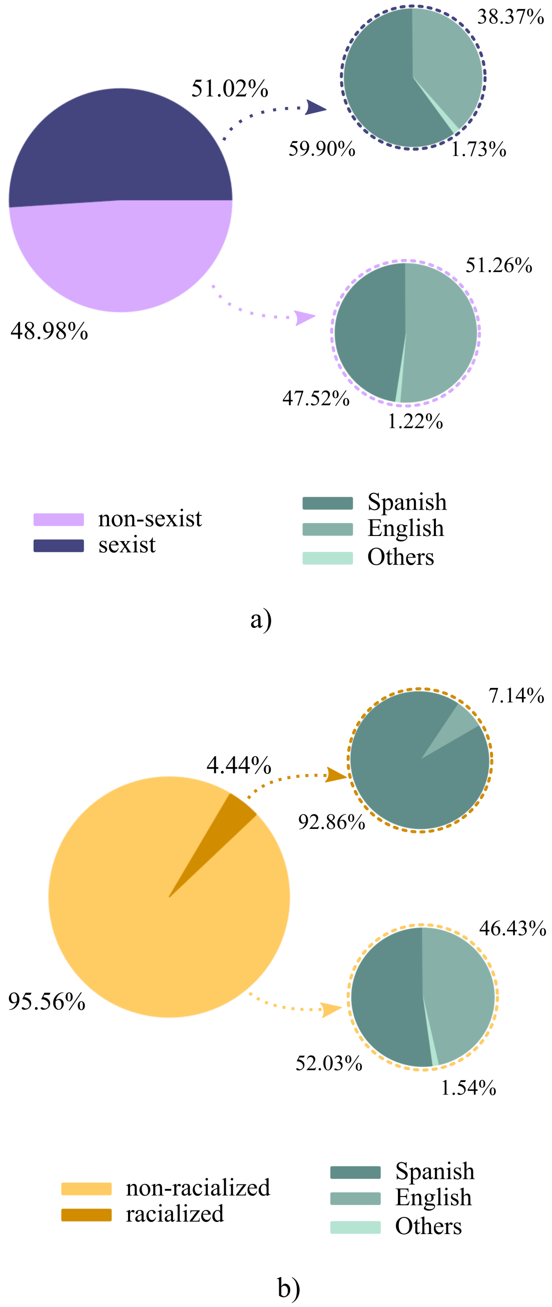


FIG. 5. a) Percentage of total sexist and non-sexist songs. Also, percentage of languages for each category. b) Percentage of total racialized and non-racialized songs. Also, percentage of languages for each category.

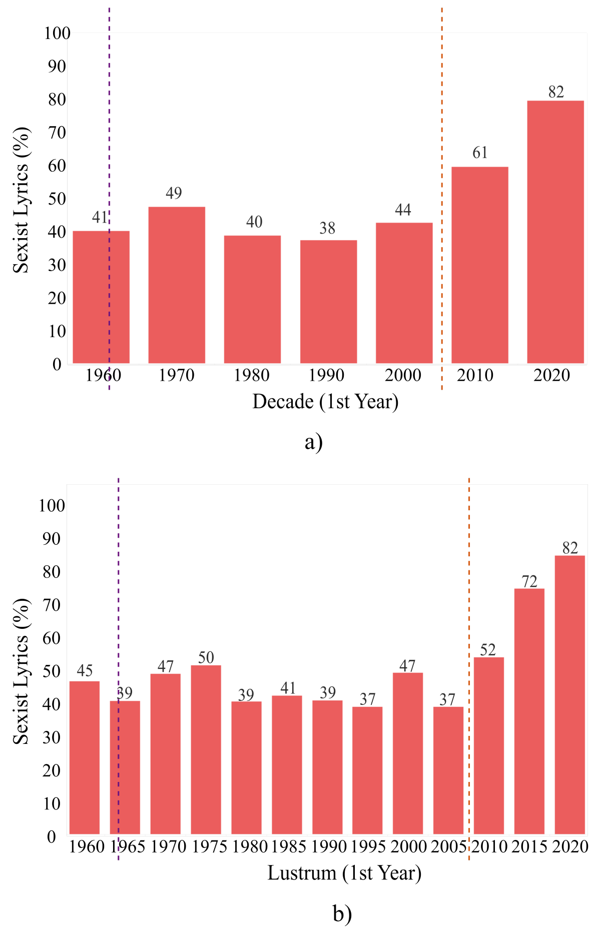


FIG. 6. Top: percentage of sexist songs per decade. Bottom: percentage of sexist songs per lustrum. The vertical dashed lines separate the three data sources used to obtain lists of popular songs: "Los Superventas," "Los 40 Principales," and Spotify.

progression of sexism in smaller groups of years to obtain a more precise analysis. Therefore, Figure 6 b) displays the evolution of sexism during the last five-year periods since 1960. Similarly to before, the proportion of sexist lyrics in the songs remains relatively constant between 37% and 50% during these periods until 2000. However, unlike in the previous examination, the trend remains consistent between 2000 and 2010. That is, between 37% and 53% of the lyrics contain sexism. With this visualization, we can more accurately determine that the primary increase in sexism in lyrics occurred during the 2015s.

Altogether, the results indicate that the most listened to songs in Spain have consistently contained sexist language throughout the years. Moreover, the prevalence of sexism in song lyrics has significantly increased in contemporary times compared to previous decades. Based on these findings, we cannot definitively conclude that there is an overall increase in sexism today. However, it is evident that the current popular music on Spotify tends to exhibit more sexist elements compared to what was popular on "Los 40 Principales" and "Los Superventas" in previous decades. This could have a more significant impact on listeners, particularly those who consume music through streaming platforms, which nowadays is the majority of people.

D. Markers of sexism

Word frequencies. We conducted a thorough analysis by comparing the frequency of words in sexist lyrics to non-sexist lyrics. To achieve this, we used scattertext (Kessler, 2017), an open-source tool that generates interactive scatterplots. In our study, we plotted the frequency of words in the non-sexist (or non-racialized) set on the x-axis and the frequency of words in the sexist (or racialized) set on the y-axis. Additionally, the interactive version displays a list next to the plot of the most frequent words for each category and the most characteristic ones overall. The data plotted consisted of the 250 most common words in the entire set. We followed the same procedure for racialized and non-racialized lyrics

We illustrate the relationship between vocabulary and sexism in Figure 7 (for English) and Figure 9 (for Spanish, in the Appendix). Despite the model being trained on paragraphs to understand the contextual usage of words, this analysis highlights a group of words that have a high probability of making a paragraph sexist. In other words, if a paragraph contains the words shown in the top-left quadrant of the graphs and avoids the words in the bottom-right quadrant, it is likely to be classified as sexist, which confirms the accuracy of the model's inferences. The pattern appears to be consistent across English and Spanish songs. The most commonly used words in sexist paragraphs revolve around pleasure, women, the body, and sensual movements. Additionally, words like "crazy" are more prevalent in sexist paragraphs than non-sexist ones. Conversely, the least frequent words in sexist paragraphs are associated with diverse states of mind, thoughts, dreams, life, and Earth places, which we could consider to be more poetic themes. Lastly, words related to love, but not explicitly sexual acts, are frequently used in both sexist and non-sexist paragraphs.

Word shifts. We also inspect the number of songs in each category of sexism within the manually labeled dataset to identify any patterns or characteristics that could aid in distinguishing between the categories.

To conduct this analysis, we used *word shifts* to measure the differences between the sets of words for each sexism category and the entire set of words in the sexist lyrics report. As described in section §III E, we employed shifterator, a generalized word shift graph that visually displays the words that contribute to the deviation between two sets of texts using a weighting measure. Specifically, we utilized proportion shifts.[9] due to their simplicity and the fact that they are easy to interpret.

In Figure 8 a), we present the most frequent words in hyper-sexualized or objectifying language compared to the whole set of words in the sexist lyrics dataset, using word shifts. Many of these words refer to the body or sexual acts and situations. For example, words such as *loco* (crazy man), *cuerpo* (body), *piel* (skin), *sabe* (tastes), *mojada* (wet woman), *culo* (butt), *duro o dura* (hard), and *tas* (you are), refer to the body and may have a sexual connotation in this context. Other words such as *beso* (kiss), *hacer* (make), *gocen* (enjoy), *gusta* (like), *cama* (bed), *gimme* (give me), and *darte* (give you), are more related to sexual acts in the context of sexist lyrics.

In the analysis of possessive and controlling language, Figure 8 b) shows the prevailing words in the entire set of sexist lyrics. We observed that words such as *dale* (give), *mueve* (move), *menea* (shake), *poompoom* (referring to a twerk movement), and *dime* (tell me) are frequent. These words are verbs in the imperative form and can be interpreted as commands given by a man to a woman to move or say certain things. Additionally, words referring to the body with sexual connotations, such as *duro* (hard), *cuerpo* (body), *culo* (butt), and *cama* (bed), are common. Furthermore, the word *loco* (crazy man), which is often related to the idea that a woman's body drives a man "crazy," also appears repeatedly.

Finally, we note that the Spanish word *chica* (girl) is not commonly used in lyrics to refer to a young woman. Instead, we observe the use of words such as the English word "girl," the English word "baby," or the diminutive *mami* (a term used to address one's mother).

E. Markers of racialization

We studied racialization adapting the method described in the previous section. In Figure 10 (in the Appendix), we plotted the most frequent words in racialized paragraphs against the non-racialized ones in the Spanish songs. As mentioned in section §IV B and displayed in Figure 5 b), the proportion of racialized songs in English is notably lower than in Spanish, making it difficult to draw any definitive inferences from the English analysis. Nevertheless, in the Spanish corpus, we detected a cluster of words that frequently appear in racialized paragraphs, but not in non-racialized ones. These words include *morena* (brunette/dark-skinned), *amiga* (female friend), *niña* (little girl), or *nena* (girl), as well as words that refer to the body, such as *labios* (lips), *dura* (hard body), or *buena* (fine, usually in the context of "*estás buena*" referring to desirable body shape). These words are also associated with women, which corroborates the analysis presented in section §IV B.

V. Conclusions and Discussion

Our main conclusions can be summarized as follows:

1. Sexism in music lyrics is highly prevalent and may have increased over time. Overall, 51% of songs in our sample contain sexist speech. While currently

[9] See more in detail here: `https://shifterator.readthedocs.io/en/latest/cookbook/frequency_shifts.html`

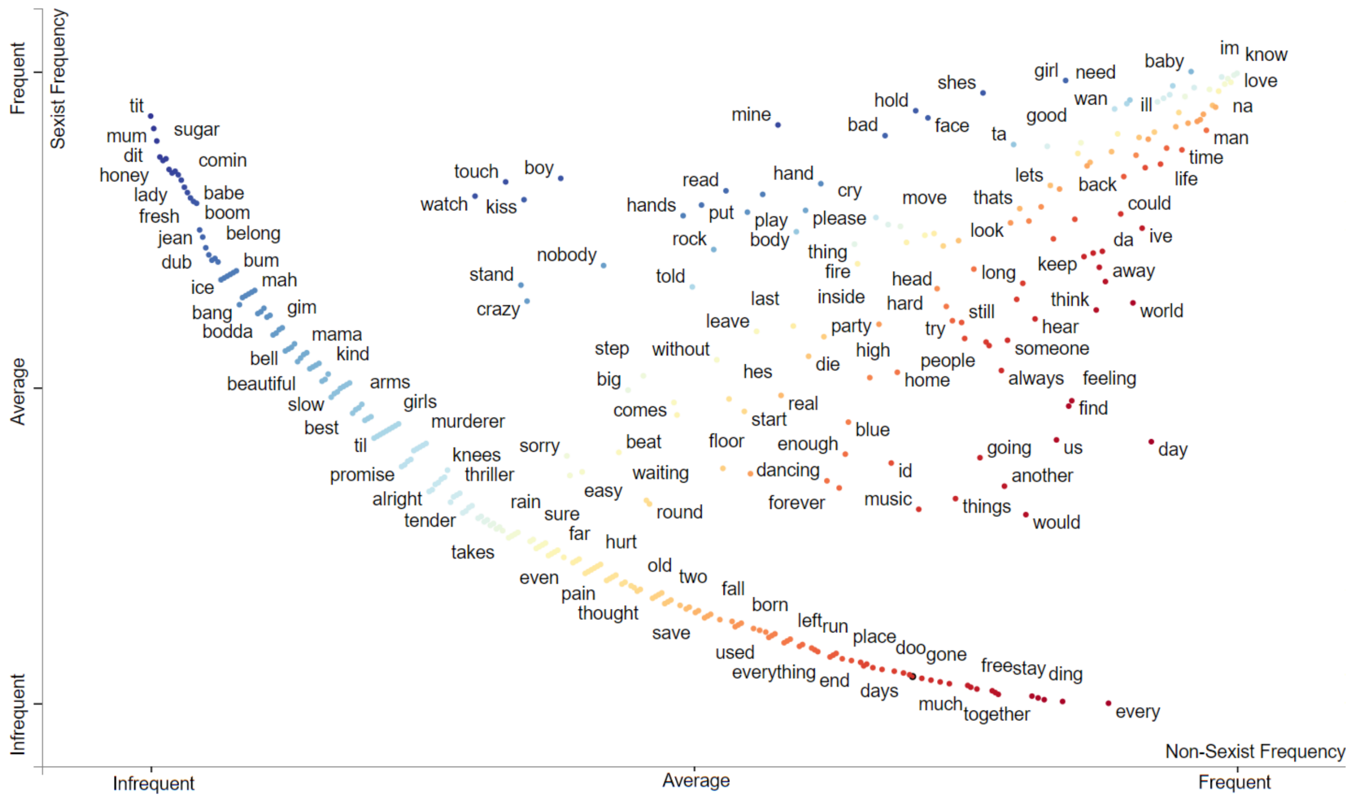


FIG. 7. Most frequent words in sexist and non-sexist paragraphs in English. On the top left, colored in blue, there are the most frequent words in sexist paragraphs that are not frequent at all in non-sexist paragraphs. On the bottom right there are the least frequent words in sexist paragraphs that are hugely frequent in non-sexist paragraphs. Finally, in the top right, there are the words that are very frequent in both sexist and non-sexist paragraphs.

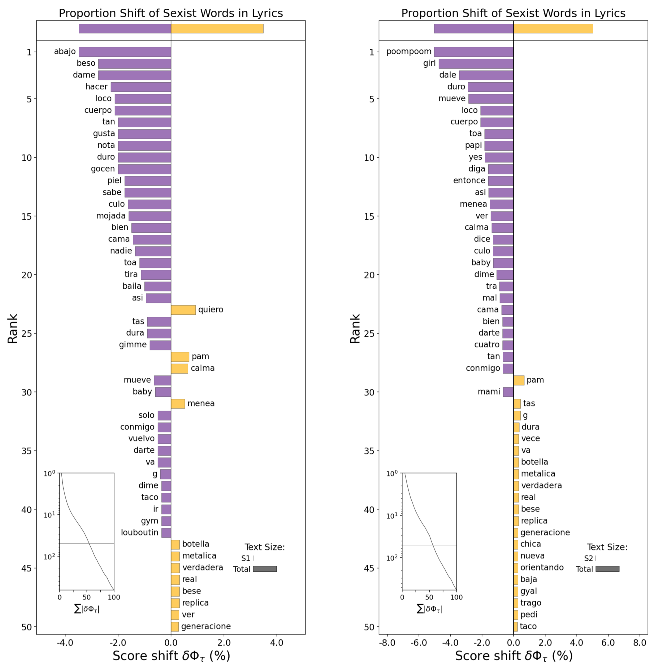
Spain scores relatively well in terms of Gender Equality Index, compared to other EU countries[10], historically sexism has been an issue (ABC, 2010, Solsten and Meditz, 1988) and we could say that this is reflected in our results. According to our findings, sexism directed towards women in the most popular songs in Spain has increased significantly over the last two decades, particularly since 2015. This may have happened simultaneously with the fact that streaming platforms overtook radio, so it is challenging to separate these two factors.

Music is an artistic form of expression, and has an important role in both emotions and processes of identification and identity. Therefore, lyrics should be a tool for positive social change, not contrariwise, and provide an opportunity to fight discrimination. This is not to fall into a naive view, but to critically describe how cultural production portrays and relates to our society. The *increased* prevalence of sexism in the last decade might be disappointing considering the impact that feminist movements have had recently in Spain, particularly the massive demonstrations for International Women's Day in 2018, 2019, and 2020. The music lyrics we have analyzed do not show a less sexist society; on the contrary, they suggest that the Spanish society still has a vast amount of sexism in it. Again, this might be because of the shift from radio to streaming, but we remark that streaming platforms are less likely to impose editorial controls that are as strict as the ones imposed by radio stations decades ago.

2. Some categories of sexist speech are more prevalent in lyrics than others. Our observations indicate that the most common categories of sexism in music lyrics are related to controlling and possessive behavior and hypersexualization. Consistently, prevalent words used in sexist-tagged songs revolve around pleasure, women, the body, and sensual movements. Additionally, the most frequent terms used for controlling and possessing behavior are imperative verbs, while for hypersexualization and objectification of women the most commonly recurring words pertain to sexual acts or circumstances.

3. Racialized speech is concentrated in a few songs. We estimate that at least 4% of the songs contain

[10] `https://eige.europa.eu/publications/gender-equality-index-2021-health`



a) b)

FIG. 8. Word shifts (shifterator) of *hyper-sexualization and women objectification* (S1) and *control and possession behavior* (S2). a) *Hyper-sexualization and women objectification* most characteristic words compared to sexist characteristic words but non-characteristic of this category. b) *Control and possession behavior* most characteristic words compared to sexist characteristic words but non-characteristic of this category.

racialized references, and while this number is relatively small, most of them are also sexist. This is constrained by the limitations we describe in the next section, particularly the low prevalence of racialized text in our training set. Certainly, it does not imply that there is no racialization at all. Actually, if we could compare lyrics with music videos (Álvarez-Cueva et al., 2021), we would probably see that there is more racialization in the images. Moreover, even though the percentage of racialized songs was low, the majority of them were also sexist. Additionally, we noticed that popular songs in Spain that are sexist and racialized are more likely to be in Spanish than in English, while the opposite is true for non-sexist and non-racialized songs. We critically argue that this result can be discussed from a gender perspective, considering the continued exoticization of Latina, and to some extent, Hispanic women in general. In other words, the media still portrays them as "exotic," not only in terms of cultural portrayals but also with respect to sexist attitudes (Arrizón, 2008).

In conclusion, as our analysis confirmed, sexism is presently a huge matter of concern in Spain. It implicates society at many layers and is clearly visible in music lyrics. This points out that sexism is systemic, and ending sexism requires the implication of various societal actors at many levels.

VI. Limitations and Future Research

First, since this study is the first attempt to gather sexism and racialization in a study focused on contemporary music in Spain, the codification represented a challenge. Available datasets containing sexist speech were extracted from entirely different domains (such as social media) and were neither sufficient nor adequate for our purposes. While most lyrics were labelled/reviewed by two coders, authors of this study, we also employed crowdsourcing during labeling to mitigate this problem. However, a broader sexism labeling exercise would greatly contribute to performing more accurate categorization of lyrics, and thus obtaining more precise statistics.

Second, before training any model with a dataset, data cleaning is usually performed. Lyrics are usually written in an informal way, and transcriptions are normally user-generated and unofficial. Songs are often made to sing and not to read. Thus, they are made of spoken vocabulary and its text is commonly not grammatically correct, which makes them harder to read. In addition, there are elements that appear in the lyrics such as sounds or onomatopoeia that are hard to parse and interpret. Therefore, during the study we find some informal words and expressions which sometimes hamper the procedure and the analysis.

Third, categorizing a song as sexist based on the presence of sexism in a single labeled paragraph may not be the most accurate approach as it disregards the broader context. Given the significant advancements in natural language processing (NLP) today, it would be desirable (but challenging) to develop a classifier trained on complete songs to retain the necessary context and provide more nuanced analysis. More generally, while automatic annotation of text allows us to process a quantity of text that we would not be able to annotate manually given our resources, as said earlier, we recognize that this process is sensitive to some extent to noise and bias, while providing a solid starting point for further exploration.

Fourth, the automatization of the lyrics gathering part could be improved to avoid loosing data and save time when doing things manually. We were not able to download about 2.7% of the songs automatically, possibly due to differences in the way their title was written, or to text coding errors. However, if we look by decade, the ones with fewer failures to download music lyrics were the last two, which means that our results would not be substantially different as these decades are also the ones where sexism was more observed.

Fifth, the small prevalence of racialized text poses challenges to automated detection, which requires large and varied training sets. A systematic effort to obtain a broader set reflecting racialization in music lyrics could alleviate this problem.

Sixth, gender bias in word embeddings is well documented and may affect our classifier results. There is an active research direction exploring how to mitigate this impact effect (Bolukbasi et al., 2016, Brunet et al., 2019, Dev and Phillips, 2019, Schmahl et al., 2020, Swinger et al., 2019).

Acknowledgements

(To be added in camera-ready version.)

ABC (2010, Sep). Muerte por adulterio en españa, un "derecho" del marido hasta 1963.

Alvarez-Cueva, P. (2022). Music to face the lockdown: An analysis of covid-19 music narratives on individual and social well-being. *Social Inclusion 10*(2), 6–18.

Álvarez-Cueva, P., M. Figueras Maz, and P. Medina Bravo (2021). Evolución de la heteronormatividad a partir de una categorización de los estereotipos de género: análisis de los videoclips musicales más populares. *El Profesional de la información. 2021; 30 (5): e300501.*.

Anderson, C. A., N. L. Carnagey, and J. Eubanks (2003). Exposure to violent media: the effects of songs with violent lyrics on aggressive thoughts and feelings. *Journal of personality and social psychology 84*(5), 960.

Ángeles Díaz, F. E. (2019). Estereotipos, roles de género y expresiones de violencia en relación a la música: el caso del reggaetón entre jóvenes limeños.

Ara, A. and R. Gopalakrishna (2020). A study on emotion identification from music lyrics. In *International Conference of Reliable Information and Communication Technology*, pp. 396–406. Springer.

Araújo Lima, R. d., R. C. C. d. Sousa, H. Lopes, and S. D. J. Barbosa (2020). Brazilian lyrics-based music genre classification using a blstm network. In *International Conference on Artificial Intelligence and Soft Computing*, pp. 525–534. Springer.

Arrizón, A. (2008). Latina subjectivity, sexuality and sensuality. *Women & Performance: a journal of feminist theory 18*(3), 189–198.

Artetxe, M. and H. Schwenk (2019). Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond. *Transactions of the Association for Computational Linguistics 7*, 597–610.

Attwood, F. (2009). The sexualization of western culture.

Aubrey, J. S., H. Gamble, and R. Hahn (2017). Empowered sexual objects? the priming influence of self-sexualization on thoughts and beliefs related to gender, sex, and power. *Western Journal of Communication 81*(3), 362–384.

Barman, M. P., A. Awekar, and S. Kothari (2019). Decoding the style and bias of song lyrics. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1165–1168.

Bearman, S. and M. Amrhein (2014). Girls, women, and internalized sexism. *Internalized oppression: The psychology of marginalized groups*, 191–225.

Bennett, A. (2001). *Cultures of popular music*. McGraw-Hill Education (UK).

Bolukbasi, T., K.-W. Chang, J. Y. Zou, V. Saligrama, and A. T. Kalai (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems 29*.

Brunet, M.-E., C. Alkalay-Houlihan, A. Anderson, and R. Zemel (2019). Understanding the origins of bias in word embeddings. In *International conference on machine learning*, pp. 803–811. PMLR.

Bryant, C. D. (1982). *Sexual deviancy and social proscription: The social context of carnal behavior*. Human Sciences Press.

Butt, S., N. Ashraf, G. Sidorov, and A. Gelbukh (2021). Sexism identification using bert and data augmentation-exist2021. In *International Conference of the Spanish Society for Natural Language Processing SEPLN 2021, IberLEF 2021*.

Cadena Ruiz, D. E. and J. M. Ortiz Lema (2018). Violencia de género en la música del discurso en la letra de las canciones del reggaetón. B.S. thesis, Ecuador, Latacunga: Universidad Técnica de Cotopaxi UTC.

Caliskan, A., J. J. Bryson, and A. Narayanan (2017). Semantics derived automatically from language corpora contain human-like biases. *Science 356*(6334), 183–186.

Capodilupo, C. M. (2017). Institutional sexism. *The SAGE encyclopedia of psychology and gender*, 941–942.

Charlesworth, T. E., V. Yang, T. C. Mann, B. Kurdi, and M. R. Banaji (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across child and adult language corpora of more than 65 million words. *Psychological Science 32*(2), 218–240.

Church, K. W. (2017). Word2vec. *Natural Language Engineering 23*(1), 155–162.

Cortes, C. and V. Vapnik (1995). Support-vector networks. *Machine learning 20*(3), 273–297.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *u. Chi. Legal f.*, 139–168.

Daykin, N. de viggiani, n., moriarty, y., & pilkington, p.(2017). music-making for health and wellbeing in youth justice settings: Mediated affordances and the impact of context and social relations. *Sociology of Health & Illness 39*(6), 941–958.

Dev, S. and J. Phillips (2019). Attenuating bias in word vectors. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 879–887. PMLR.

Devlin, J. and M.-W. Chang (2018). Open sourcing bert: State-of-the-art pre-training for natural language processing. *Google AI Blog 2*.

Dhoest, A., R. Herreman, and M. Wasserbauer (2015). Into the groove: exploring lesbian and gay musical preferences and'lgb music'in flanders. *Observatorio (OBS*) 9*(2), 207–223.

Domènech, A. (2018, Mar). El potente mensaje feminista de 'lo malo' de aitana war (ot).

Eguizábal Jiménez, L. (2021). La radio musical española de los sesenta y su influencia en el desarrollo de la música popular.

Eisikovits, Z. and E. Buchbinder (1997). Talking violent: A phenomenological study of metaphors battering men use. *Violence Against Women 3*(5), 482–498.

Evans, A. and S. Riley (2014). *Technologies of sexiness: Sex, identity, and consumer culture.* Sexuality, Identity, and Socie.

EXIST (2021). Exist: sexism identification in social networks.

Eze, S. U. (2020). Sexism and power play in the nigerian contemporary hip hop culture: The music of wizkid. *Contemporary music review 39*(1), 167–185.

Flynn, M. A., C. M. Craig, C. N. Anderson, and K. J. Holody (2016). Objectification in popular music lyrics: An examination of gender and genre differences. *Sex roles 75*(3), 164–176.

Frenda, S., B. Ghanem, M. Montes-y Gómez, and P. Rosso (2019). Online hate speech against women: Automatic identification of misogyny and sexism on twitter. *Journal of Intelligent & Fuzzy Systems 36*(5), 4743–4752.

Gallagher, R. J., M. R. Frank, L. Mitchell, A. J. Schwartz, A. J. Reagan, C. M. Danforth, and P. S. Dodds (2021). Generalized word shift graphs: a method for visualizing and explaining pairwise comparisons between texts. *EPJ Data Science 10*(1), 4.

Garg, N., L. Schiebinger, D. Jurafsky, and J. Zou (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences 115*(16), E3635–E3644.

Genius Media Group (2021). Genius api. `https://docs.genius.com/`. Accessed on: 2023-02-07.

Gill, R. (2003). From sexual objectification to sexual subjectification: The resexualisation of women's bodies in the media. *Feminist media studies 3*(1), 100–106.

Gill, R. (2011). Sexism reloaded, or, it's time to get angry again! *Feminist media studies 11*(01), 61–71.

Gill, R. (2017). The affective, cultural and psychic life of postfeminism: A postfeminist sensibility 10 years on. *European journal of cultural studies 20*(6), 606–626.

Gillis, N. B. (2021). Sexism in the judiciary. *arXiv preprint arXiv:2106.15103*.

Good, J. J. (2017). Benevolent sexism. *The SAGE encyclopedia of psychology and gender*, 136–137.

Greitemeyer, T. (2009). Effects of songs with prosocial lyrics on prosocial thoughts, affect, and behavior. *Journal of experimental social psychology 45*(1), 186–190.

Grosz, D. and P. Conde-Cespedes (2020). Automatic detection of sexist statements commonly used at the workplace. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 104–115. Springer.

Grubbs, J. B. (2017). Ambivalent sexism. *The SAGE encyclopedia of psychology and gender*, 61–63.

Guarinos, V. and A. M. S. Valdellós (2020). El videoclip social. análisis de su narrativa y estética. *methaodos. revista de ciencias sociales 8*(1), 120–129.

Guest, E. (2021). online-misogyny-eacl2021. `https://github.com/ellamguest/online-misogyny-eacl2021/tree/main/data`.

Gupta, A., H. Ladia, S. Bansal, and A. Arora (2021). Brave men and emotional women: Analyzing gender bias in bollywood songs. In *Proceedings of International Conference on Computational Intelligence and Data Engineering*, pp. 143–153. Springer.

Hack, T. (2017). Hostile sexism. *The SAGE encyclopedia of psychology and gender*, 895–896.

Hall, S. (1992). *The question of cultural identity: modernity and its futures.*

Hall, S. (2011). The neo-liberal revolution. *Cultural studies 25*(6), 705–728.

Hanley, J. A. and B. J. McNeil (1982). The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology 143*(1), 29–36.

Illescas, J. E. (2015). La dictadura del videoclip. *El viejo topo* (335), 32–38.

Jha, A. and R. Mamidi (2017). When does a compliment become sexist? analysis and classification of ambivalent sexism using twitter data. In *Proceedings of the second workshop on NLP and computational social science*, pp. 7–16.

Joulin, A., E. Grave, P. Bojanowski, and T. Mikolov (2016). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Kessler, J. S. (2017). Scattertext: a browser-based tool for visualizing how corpora differ. *arXiv preprint arXiv:1703.00565*.

Kim, Y., Y.-I. Chiu, K. Hanaki, D. Hegde, and S. Petrov (2014). Temporal analysis of language through neural language models. *arXiv preprint arXiv:1405.3515*.

Koskei, J. C., M. Barasa, and B. Manyasi (2018). Stereotypical portrayal of women in kipsigis secular songs.

Kotarba, J. A. (1994). The postmodernization of rock and roll music: The case of metallica. *Adolescents and their Music: If it's too Loud, you're too Old, New York, NY, Garland Publishing*, 141–163.

Kozlowski, A. C., M. Taddy, and J. A. Evans (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review 84*(5), 905–949.

LOS40 (2018, Feb). Lista 40 principales 1966.

Mahedero, J. P., Á. MartÍnez, P. Cano, M. Koppenberger, and F. Gouyon (2005). Natural language processing of lyrics. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pp. 475–478.

Meinecke, C. and S. Jänicke (2021). Detecting text reuse and similarities between artists in rap music through visualization.

Nava-Reyes, M. A., J. L. Rojas-Solís, L. M. G. Amador, and L. A. M. Quintero (2018). Gender roles, sexism and myths of romantic love in mexican adolescents. *Revista Interamericana de Psicología 52*(1).

Oyeka, C. N. (2019). Sexism in lyrics. *Culture, Precepts, and Social Change in Southeastern Nigeria: Understanding the Igbo*, 141.

Papadaki, E. L. (2010). Feminist perspectives on objectification.

Parikh, P., H. Abburi, P. Badjatiya, R. Krishnan, N. Chhaya, M. Gupta, and V. Varma (2019). Multi-label categorization of accounts of sexism using a neural framework. *arXiv preprint arXiv:1910.04602*.

Pennington, J., R. Socher, and C. D. Manning (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543.

Pintado, N. R. (2017, Nov). Discos superventas en españa 1940 – 2017.

Rasmussen, E. E. and R. L. Densley (2017). Girl in a country song: Gender roles and objectification of women in popular country music across 1990 to 2014. *Sex Roles 76*(3), 188–201.

Richardson, L. (1991). Postmodern social theory: Representational practices. *Sociological theory 9*(2), 173–179.

Rodriguez, F. (2020). Metwo: Machismo and sexism twitter identification dataset. `https://github.com/franciscorodriguez92/MeTwo`.

Rospocher, M. (2021). Explicit song lyrics detection with subword-enriched word embeddings. *Expert Systems with Applications 163*, 113749.

Rospocher, M. (2022). Detecting explicit lyrics: a case study in italian music. *Language Resources and Evaluation*, 1–19.

Rudolph, M. and D. Blei (2018). Dynamic embeddings for language evolution. In *Proceedings of the 2018 World Wide Web Conference*, pp. 1003–1011.

Samghabadi, N. S., P. Patwa, S. Pykl, P. Mukherjee, A. Das, and T. Solorio (2020). Aggression and misogyny detection using bert: A multi-task approach. In *Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying*, pp. 126–131.

Schaefer, R. T. (2008). *Encyclopedia of race, ethnicity, and society*, Volume 1. Sage.

Schmahl, K. G., T. J. Viering, S. Makrodimitris, A. N. Jahfari, D. Tax, and M. Loog (2020). Is wikipedia succeeding in reducing gender bias? assessing changes in gender bias in wikipedia using word embeddings. In *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*, pp. 94–103.

Schneider, C. J. (2011). Culture, rap music,"bitch," and the development of the censorship frame. *American behavioral scientist 55*(1), 36–56.

Sharifirad, S. and S. Matwin (2019). When a tweet is actually sexist. a more comprehensive classification of different online harassment categories and the challenges in nlp. *arXiv preprint arXiv:1902.10584*.

Skeggs, B. (1997). Formations of class & gender: Becoming respectable. *Formations of Class & Gender*.

Slim, K. (2019). Feminism vs. sexism in lyrics: A portrait of women in recent music. `https://github.com/axnyang/CS401/blob/master/data/train_lyrics.csv`.

Smiler, A. P., J. W. Shewmaker, and B. Hearon (2017a). From "i want to hold your hand" to "promiscuous": Sexual stereotypes in popular music lyrics, 1960–2008. *Sexuality & Culture 21*(4), 1083–1105.

Smiler, A. P., J. W. Shewmaker, and B. Hearon (2017b). From "i want to hold your hand" to "promiscuous": Sexual stereotypes in popular music lyrics, 1960–2008. *Sexuality & Culture 21*(4), 1083–1105.

Sollee, K. (2015). Mining the motherload: Mastodon's# twerkgate and sexual objectification in metal. *Metal Music Studies 1*(2), 271–277.

Solsten, E. and S. Meditz (1988). Social values and attitudes. *Spain: A Country Study*.

Swidler, A. (1996). Newsletter of the sociology of culture. *Culture*.

Swinger, N., M. De-Arteaga, N. T. Heffernan IV, M. D. Leiserson, and A. T. Kalai (2019). What are the biases in my word embedding? In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 305–311.

U.N.O.D.C. (2018, Oct). Integrity ethics module 9 key issues: Forms of gender discrimination.

Van Bohemen, S., L. Den Hertog, and L. van Zoonen (2018). Music as a resource for the sexual self: An exploration of how young people in the netherlands use music for good sex. *Poetics 66*, 19–29.

Van Maanen, J. (1995). Representation in ethnography. *SAGE Publications Inc*.

Van Rijsbergen, C. (1983). Information retrieval.

Vandenbosch, L., D. Vervloessem, and S. Eggermont (2013). "I might get your heart racing in my skin-tight jeans": Sexualization on music entertainment television. *Communication Studies 64*(2), 178–194.

Walker, S. (1994). *Hate speech: The history of an American controversy.* U of Nebraska Press.

Walter, N. (2010). *Living dolls: The return of sexism.* Hachette UK.

Ware, S. J. (2020). Paternalism or gender-neutrality? *Conn. L. Rev. 52*, 537.

Wex Definitions Team (2020, Jun). Gender bias.

Whiteley, S. (1992). On record: Rock, pop and the written word. edited by simon frith and andrew goodwin. london: Routledge, 1990. 492 pp. *Popular Music 11*(1), 123–126.

Wright, C. and E. Qureshi (2015). The relationship between sexual content in music and dating and sexual behaviors of emerging adults. *Howard Journal of Communications 26*(3), 227–253.

Appendix

TABLE III. Titles with random sample of songs with 100% of paragraphs marked as sexist (shaded rows are songs by female artists)

Title	Artist	% Racialized	Decade	Genre
Twist and shout	The Beatles	0	1960	Rock
Quince años tiene mi amor	Duo Dinamico	0	1960	Pop
Balada gitana	Duo Dinamico	50	1960	Pop
Tonight my love tonight	Paul Anka	0	1960	Pop
Delilah	Tom Jones	0	1960	Pop
Gorrion	Miguel Gallardo	0	1970	Pop
Can the can	Suzi Quatro	0	1970	Rock
Brother Louie	Modern Talking	0	1980	Pop
Somethig bout you baby I like	Status Quo	0	1980	Pop
A woman loves a man	Joe Cocker	0	1980	Country
Maria	Blondie	0	1990	Pop
Bumpy Ride	Mohombi	0	2010	Dance-Pop
Rude Boy	Rihanna	0	2010	Dance
Cosas locas	Danny Romero	0	2010	House
Sugar	Maroon 5	0	2010	Pop
Gyal you a party animal	Charly Black	0	2010	Dance
Traicionera	Sebastian Yatra	0	2010	Urban Latin
Unica	Ozuna	50	2010	Urban Latin
La Rubia	La Nueva Escuela	100	2020	Urban Latin
Relacion	Sech	0	2020	Urban Latin
Yo perreo sola	Bad Bunny	50	2020	Urban Latin
PAM	Justin Quiles	0	2020	Urban Latin
Flamenkito	Lerica	20	2020	Pop
Rutina	Myke Towers	0	2020	Urban Latin
Diosa	Myke Towers	40	2020	Urban Latin
Explicito	Myke Towers	0	2020	Urban Latin
Hola	Dalex	0	2020	Urban Latin
La dificil	Bad Bunny	0	2020	Urban Latin
Loco	Beéle	0	2020	Pop
Indeciso	Reik	0	2020	Urban Latin
China	Annuel AA	0	2020	Urban Latin
Kesi	Camilo	0	2020	Urban Latin
Tutu	Camilo	0	2020	Urban Latin
Travesuras	Nio Garcia	0	2020	Urban Latin
A donde vamos	Morat	0	2020	Pop
Nathy Peluso Bzrp Music Sessions Vol. 36	Bizarrap	0	2020	Urban Latin
Reloj	Rauw Alejandro	0	2020	Urban Latin
Veneno	Delaossa	0	2020	Urban Latin

TABLE IV. Titles with 41 random non-sexist songs (shaded rows are songs by female artists)

Title	Artist	Decade	Genre
Si je chante	Sylvie Vartan	1960	Pop
Soy rebelde	Jeanette	1970	Latin ballad
Melina	Camilo Sesto	1970	Latin ballad
Stayin alive	Bee gees	1970	Disco
Algo mas	Camilo Sesto	1970	Pop
Esta noche me quiero descolgar	Trupita	1980	Techno
Breakthoven	Baron Rojo	1980	Heavy Metal
Never gonna give you up	Rick Astley	1980	Dance-pop
Old before I die	Robbie Williams	1990	Rock
Pump up the jam	Technotronic	1990	House
Go west	Pet Shop Boys	1990	Disco
Have you seen her	Mc Hammer	1990	Hip-hop
Baby I love you	Big Mountain	1990	Reggae
La deuda de la mentira	Danza Invisible	1990	Rock/Pop
The man who sold the world	Nirvana	1990	Glam rock
Fields of Gold	Sting	1990	Rock
Don't speak	No doubt	1990	Pop
Dalai lama	Mecano	1990	Pop
Chihuahua	Dj bobo	2000	Eurodance
Quiero ser	Amaia Montero	2000	Pop
You're beautiful	James Blunt	2000	Rock/Pop
Aunque no te pueda ver	Alex Ubago	2000	Pop
Musica para una boda	Nacho Cano	2000	Pop
Te busqué	Nelly Furtado	2000	Reggae
Dream on	Depeche Mode	2000	Electronic
Whenever, whenever	Shakira	2000	Pop
Flames	David Guetta	2010	Electropop
It's time	Imagine Dragons	2010	Rock/Pop
Mi verdad	Maná and Shakira	2010	Rock/Pop
Reality	Lost frequencies and Janieck Devy	2010	Electronic
Terriblemente Cruel	Leiva	2010	Indie/Rock
Antes de que cuente diez	Fito y Fitipaldis	2010	Rock
Live your life	Mika	2010	Hip-hop
Roar	Katy Perry	2010	Pop
Sing	Ed Sheeran	2010	R&B
Thunder	Imagine Dragons	2010	Electropop
Wake me up	Avicii	2010	House
Malibu	Miley Cyrus	2010	Electropop
Ain't nobody loves me better	Felix Jaehn	2010	Electronic
Con la mano levanta'	Macaco and Estopa	2010	Latin pop
Where have you been	Rihanna	2010	Dance-pop

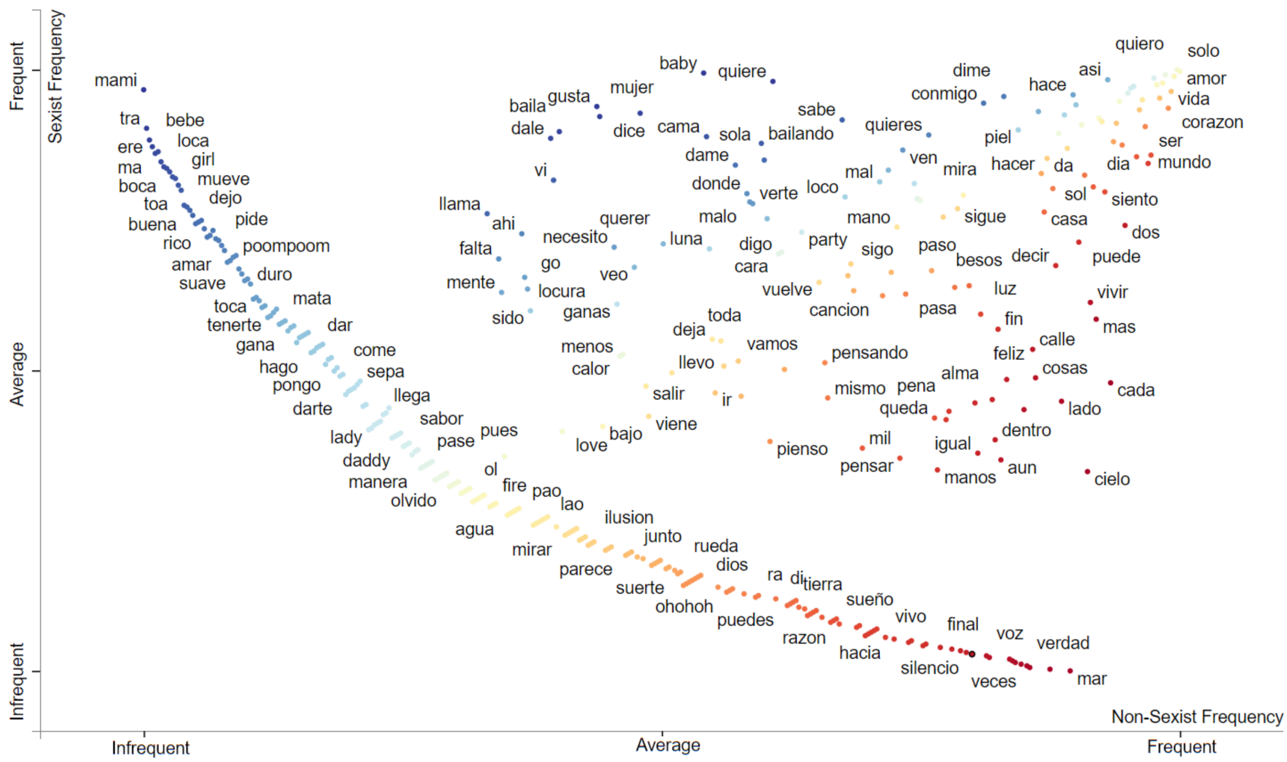


FIG. 9. Most frequent words in sexist and non-sexist paragraphs in Spanish. On the top left, colored in blue, there are the most frequent words in sexist paragraphs that are not frequent at all in non-sexist paragraphs. On the bottom right there are the least frequent words in sexist paragraphs that are hugely frequent in non-sexist paragraphs. Finally, in the top right, there are the words that are very frequent in both sexist and non-sexist paragraphs.

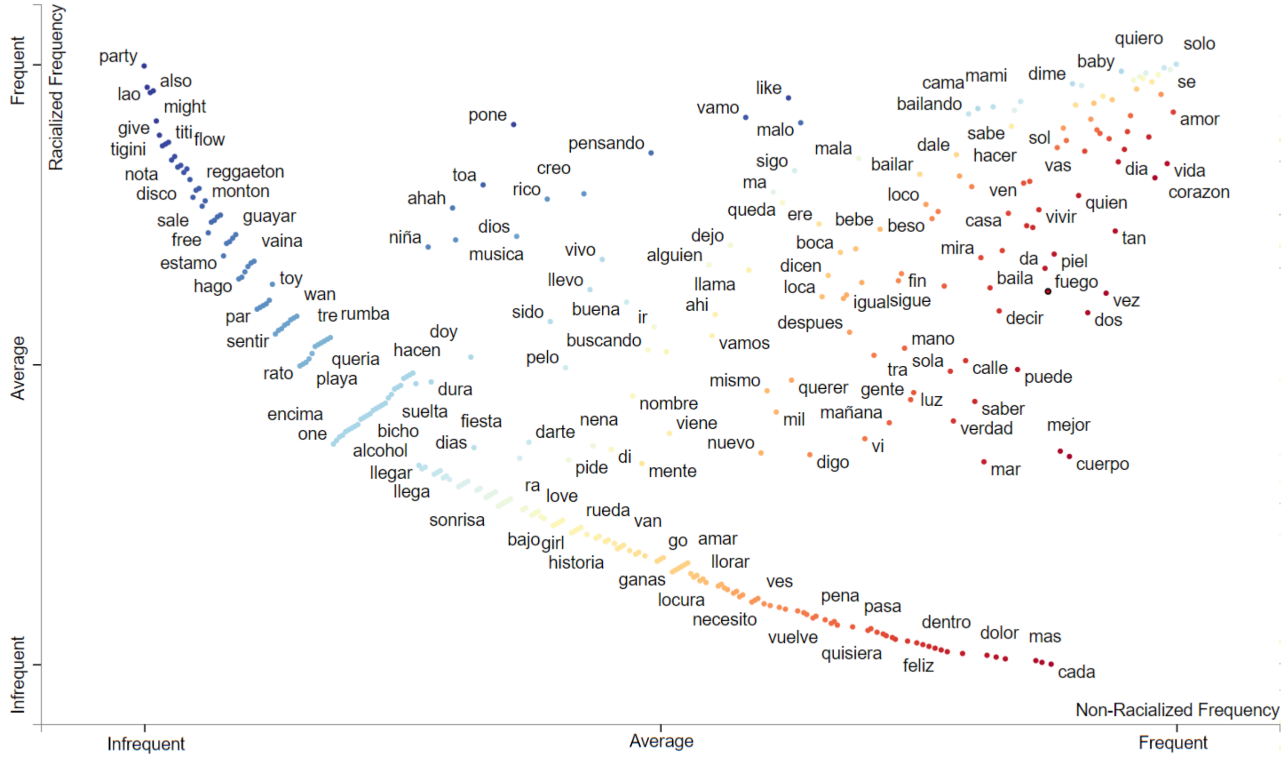


FIG. 10. Most frequent words in racialized and non-racialized paragraphs in Spanish. On the top left, colored in blue, there are the most frequent words in racialized paragraphs that are not frequent at all in non-racialized paragraphs. On the bottom right there are the least frequent words in racialized paragraphs that are hugely frequent in non-racialized paragraphs. Finally, in the top right, there are the words that are very frequent in both racialized and non-racialized paragraphs.