social sciences

MDPI

Article

Online Hate Speech and Immigration Acceptance: A Study of Spanish Provinces

Patricia Sánchez-Holgado, Javier J. Amores and David Blanco-Herrero *

Department of Sociology and Communication, University of Salamanca, 37008 Salamanca, Spain
* Correspondence: david.blanco.herrero@usal.es

Abstract: Online hate speech against migrants and refugees poses a grave challenge for coexistence and democracy. However, it also offers an opportunity to measure social acceptance towards this group. Using the Intergroup Contact and the Mediated Intergroup Contact Theory, and an already validated methodology, this article seeks to validate whether the use of hate speech as a predictor of social acceptance is valid at a provincial level in Spain. Contrasting 97,710 tweets about migrants and refugees with secondary data from public Spanish institutions about acceptance of immigration and foreign population, no correlation was observed, rejecting the main hypotheses, and hinting that the application of this approach might not be recommended for smaller entities, such as provinces (NUTS 3). However, the study offers descriptive data about racist hate speech spread on Twitter in Spain, and also discusses the need for more studies using big data to increase knowledge about online hate speech against migrants and refugees.

Keywords: hate speech; racism and xenophobia; attitudes toward migration; big data; Twitter; immigrants; refugees

Citation: Sánchez-Holgado, Patricia, Javier J. Amores, and David Blanco-Herrero. 2022. Online Hate Speech and Immigration Acceptance: A Study of Spanish Provinces. *Social Sciences* 11: 515. https://doi.org/10.3390/socsci11110515

Academic Editors: Andreu Casero-Ripolles and Tuba Bircan

Received: 30 July 2022
Accepted: 10 November 2022
Published: 14 November 2022

1. Introduction

Even though it is not only a digital problem, online hate speech has also become one of the most concerning threats in modern societies. These types of discourse, when aimed at vulnerable groups, do not only harm coexistence and risk integration, but they can also lead to cases of real physical violence (Müller and Schwarz 2020). This is particularly dangerous because, as the data recorded by the Hate Crime Reporting of the Organization for Security and Co-operation in Europe (n.d.) show, hate crimes have been increasing in the last years in most countries.

Among these crimes, one of the most common causes is racism and xenophobia, which is the type of hatred behind the rejection towards migrants, refugees and foreigners in general. In this regard, and in comparative terms, we can see that Spain could be considered one of the most tolerant countries for foreigners—see, for instance, the Special Eurobarometers 437 (European Commission 2015), 469 (European Commission 2018), 493 (European Commission 2019) or 519 (European Commission 2022). However, in the last years there has been a relevant increase of anti-immigration narratives, which had been less present until not long ago (Turnbull-Dugarte 2019; Ferreira 2019). This happened mostly due to the gains of a nationalist political party such as Vox, which caused an increase in the importance of migration in the public agenda in Spain (Castromil et al. 2020), and whose discourse around migration are usually connected with hate speech (Arcila-Calderón et al. 2020b).

However, the definition of hate speech is not always easy, as there are different approaches, including legal and ethical, such as the distinction made between hate speech and offensive language made by Davidson et al. (2017), and the different levels of graveness discussed by Miró Llinares (2016). In spite of these unresolved discussions, there is a general agreement around the validity of the definition proposed by different European institutions, such as the Recommendation 97 of the Council of Europe (1997), stating that hate speech

refers to "all forms of expression which incite racial hatred, xenophobia, anti-Semitism and all forms of intolerance". This could be complemented by the Framework Decision 2008/913/JHA of 28 November 2008 of the Council of the European Union (2008), defining hate speech as "publicly inciting to violence or hatred directed against a group of persons or a member of such a group defined by reference to race, color, religion, descent or national or ethnic origin".

Using these definitions as a reference, the present article seeks to contribute to the understanding and modelling of hate speech in Spain. This text tries to fill in the existing knowledge gap regarding the small region or province level (NUTS 3 in the Nomenclature of Territorial Units for Statistics system), as most analyses tend to focus on national comparisons or, at best, the regional level (NUTS 2). This was the case of the work of Arcila-Calderón et al. (2022b), which offers a European-wide analysis; our work will seek to complement and further develop this research, paying special attention to the Spanish case. Furthermore, this study uses geolocated tweets, something that has not been common in previous works using Twitter message as their object of study, thus improving the quality of the analysis, as the geographical origin of the messages can be ensured.

Research Approach and Hypotheses

In the context depicted in the previous paragraphs, hate speech becomes not only a dangerous phenomenon to be tackled, but also a potential tool to measure social acceptance. In fact, even though the definition of online hate speech and its measure are not always clear and easy, previous studies have supported that measuring hate speech could be used as a way to evaluate attitudes towards migrants and refugees (Arcila-Calderón et al. 2020a), and in general, social acceptance towards this group (Arcila-Calderón et al. 2022b). This is important because the decrease in social acceptance towards minorities or vulnerable groups—not only foreigners, also religious minorities or members of the LGBTQ+ community, among others—is a risk for democratic and diverse societies.

Several theories support this approach. First of all, rejection is partly explained by the theory of social identity (Tajfel 1978), according to which migrants and refugees would be associated with the 'exogroup', which also leads to lower levels of acceptance (Esses et al. 2005). Furthermore, the Intergroup Contact Theory (ICT) of Allport (1954) argues that a higher intergroup contact with members of the 'exogroup'—here, foreigners—can lead to higher levels of acceptance and tolerance and more positive attitudes towards outgroup members (Abrams and Hogg 2017). Thus, our first hypothesis is:

Hypothesis H1. *Spanish provinces with a higher immigrant population will show greater levels of acceptance toward immigrants.*

At the same time, given the relevance of social media in modern societies, multiple studies have focused on the presence of hate speech on social media (Bartlett et al. 2014), including analyses about anti-immigration narratives (Kreis 2017; Chaudhry 2015). Given that hate speech is based on negative attitudes, and these attitudes could be, under the ICT, influenced by the contact experienced with members of the outgroup, we present the second hypothesis:

Hypothesis H2. *Spanish provinces with a higher immigrant population will show lower levels of hate speech toward immigrants.*

Finally, keeping in mind the theory of Mediated Intergroup Contact, the presence of online hate speech, beyond its attack and direct harm to the victims, would also act as a negative contact, thus making offline attitudes and acceptance also negative. In this regard, previous studies have stated that hatred and online rejection toward immigrants could be a way to measure attitudes (Arcila-Calderón et al. 2020a). Furthermore, social surveys present several limitations when measuring rejection toward immigrants, including the potential effects of biases (Cea d'Ancona 2009), such as the social desirability one—

answering what society expects, not what I really feel, especially around topics negatively seen by society, such as racism—the large amount of time and resources consumed, the incapacity to be constantly updated and the impossibility to reach relevant sample sizes in small geographical entities, such as provinces. However, the measure of hate speech on social media offers an alternative, and given that the expression of hate speech could be considered a representation of negative attitudes, we present the third hypothesis.

Hypothesis H3. *Spanish provinces with higher levels of hate speech online will show lower levels of acceptance toward immigrants.*

2. Materials and Methods

2.1. Sample and Procedure

To analyze the level of hate speech on Twitter and its possible relationship with the proportion of migrants and with the attitude towards migration in the Spanish provinces, different sources have been used in this study. First, a sample of geolocated tweets about migrants and refugees in Spanish published from 2015 to 2020, was used, from which the level of racist and/or xenophobic hate was detected. The downloaded data from each tweet and the keywords used for the download of the tweets can be consulted as supplementary materials. The download and collection of the tweets and their metadata was conducted via the academic version of the Twitter Application Programming Interface (API) (v.2 full-archive search endpoint), which allows the publication history to be consulted, and using the Python programming language and the NLTK, Tensorflow, Keras and Numpy libraries. This download and compilation work was supported by the infrastructure of the Castilla y León Supercomputing Center (Scayle). Location filters were included in the query to access only the tweets published in the Spanish provinces, and keywords that referred to migrants, refugees and asylum seekers were employed as linguistic filters. In addition, when collecting the tweets, a function was executed to estimate their geolocation using the coordinates of the publications in the cases that were available, as well as the entities on locations named in the messages themselves. In this way, the texts could be geocoded according to the lower division (of small regions for specific diagnoses) of the Nomenclature of Statistical Territorial Units (NUTS), for which the Nominatim coordinate geocoder was used, along with data from openstreetmap, and the NUTS code identifier, NutsFinder. It should be noted that, although the NUTS 3 classification contemplates 59 localities in Spain, for the analyzes, only the 52 provinces that contemplate the Spanish official organisms, such as the INE or the CIS, were taken into account.

Although the original number of downloaded tweets reached 124,337 messages, after removing all the data in which the NUTS 3 identifier were not detected, the total sample included 97,710 tweets about migrants and refugees geolocated at the provincial level in Spain in the 6 years of analysis. Disaggregating the sample in years, there were 23,481 tweets in 2015, 27,835 in 2016, 16,862 in 2017, 19,495 in 2018, 18,811 in 2019 and 14,707 in 2020. 2016 (28.5% of the tweets) and 2015 (24.0%) were the years in which a greater number of tweets could be collected, something that could indicate that during those years the conversation on the migratory issue was greater in Spain (those were the years in which the Refugee Crisis of the Mediterranean had a greater presence in the media). Once this sample was collected, the level of anti-immigration hate present in the tweets was identified, for which the detection tool developed and validated by Vrysis et al. (2021) within the framework of the European project Preventing Hate Against Refugees and Migrants (PHARM) was used[1].

This tool is a prototype based on a deep learning architecture, specifically Recurrent Neural Networks, capable of detecting racist and/or xenophobic hate in messages about migration in Spanish, Italian or Greek, and coming from Twitter, Facebook or YouTube, as well as other online sources. The classifier developed to detect hate in Spanish was trained with 46,012 messages about migration previously classified by humans, of which 11,117 contained explicit hate. The evaluation of the detection models yielded good results: accuracy = 0.87, F-Score = 0.87. The results of the tests of previous models with different

algorithms, parameters and training corpuses, as well as of the processes of validation and improvement of the prototype for the detection in the three languages of southern Europe, are more widely reported in Arcila-Calderón et al. (2022b).

Secondly, in order to relate these data on the anti-immigration hate that is propagated on Twitter in Spain with other existing data and analyze some of the dimensions that could eventually explain this hate speech, secondary sources were used. Specifically, official Spanish sources were used, the National Institute of Statistics (INE) and the Center for Sociological Research (CIS), from which data were collected on the proportion of foreign citizens and on the attitudes towards migration at the provincial level.

2.2. Measures and Analysis

Three main variables have been used in this study: first, the average level of hate speech toward migrants and refugees propagated on Twitter in the Spanish provinces from 2015 to 2020. The value of this measure ranges from 0 to 1, with 0 being the absence of anti-immigration hate and 1 being the presence of explicit hate. The tool was able to detect hate in each of the tweets of the sample; with the intention of extracting an indicator of the level of hate that could be related to the other measures, the average value per province and year was calculated.

Second, the proportion of foreign population in the Spanish provinces. This measure was obtained from public data from the National Institute of Statistics (INE) and reflects the percentage of foreign citizens in relation to the total population in each of the 52 Spanish provinces, and in each of the 6 years analyzed (2015–2020). The values of this variable also range between 0 and 1, with 0 meaning the absence of foreigners in the province and 1 meaning that the entire population has foreign origin.

The third and last measure that has been included in this work is the attitude of Spanish citizens toward migrants and refugees. This indicator was prepared based on data available from the CIS survey "Attitudes toward migration", from 2017, so it only reflects the attitudes of that moment. This survey was used because it is the most current of those available in the country in this area, not having data from the rest of the years present in the sample. Specifically, the indicator was generated from a scale in which eight items were included, corresponding to three of the main questions of said survey. The aim was to generate an indicator that would broadly and reliably reflect the level of support for migrants and refugees by Spanish citizens in each of the Spanish provinces in 2017. Thus, after testing different combinations of items on attitudes toward migrants and refugees who had the same question and answer format, finally those who presented greater internal consistency ($\alpha = 0.913$) were included. All the selected items were formulated in the form of statements with a valence indicating a negative attitude toward migration, in which the participants showed their degree of agreement or disagreement on a Likert scale of 1 to 4, with 1 being "strongly agree", and 4 "strongly disagree". Specifically, the eight items selected to elaborate the construct were the following:

- P.14.1. The presence of immigrants decreases the quality of healthcare;
- P.14.3. Spaniards should have preference when it comes to accessing healthcare;
- P.14.4. Although they have the same income, more health aid is given to immigrants than to Spaniards;
- P.15.2. The quality of education worsens in schools where there are many children of immigrants;
- P.15.3. Spaniards should have preference when choosing their children's school;
- P.15.4. Even though they have the same income, more school aid is given to immigrants than to Spaniards;
- P.21.4. Immigrants take jobs away from Spaniards;
- P.21.5. If someone who comes to live and work here stays unemployed for a long time, he/she should be expelled from the country.

After recoding these items, considering lost all the answers coded as "8 = does not know" or "9 = does not answer", a new variable was created that reflected the mean of

the entire construct, and later it was adapted into a percentage a value that also ranged between 0 and 1, with 0 being the most negative attitude toward migration and 1 being the most positive attitude or greater support for migrants. It should be noted that this construct, besides only representing the attitude toward migration based on survey data (with the possible limitations of internal validity that this instrument entails) existing in 2017, does not reflect the attitude from some of the Spanish provinces, specifically Ávila, Soria, La Rioja and Melilla, where the CIS survey did not offer disaggregated data.

Finally, having all the standardized measures that were of interest for the study, we proceeded to explore the existing data and contrast the proposed hypotheses, for which visual, descriptive and correlational analyzes were used. Regarding the visual analyses, interactive maps were prepared in which the different variables could be compared by province and year, for which the Flourish data visualization program was used. These comparative maps, like the descriptive tables provided in the following section, represent the average values of each of the measures explained, extracted with the statistical program SPSS. This last tool was also used to carry out statistical correlation tests and scatter plots in order to explore the validity of our hypotheses.

3. Results

To test the first hypothesis (H1), in which it is stated that the proportion of foreign population in the Spanish provinces (NUTS 3) is related to the attitude of citizens toward this group, we can see in Figure 1 the proportion of foreign population in each Spanish province. It can be observed that the provinces with the highest proportion of immigrant population are: Barcelona (0.25), Girona (0.24), Alicante (0.23), Balearic Islands (0.21) and Lleida (0.21), which are all above 20% of foreign population and all located in the Mediterranean. The lowest in immigrant population can be found in Jaén (0.026) and Córdoba (0.027), Badajoz (0.030), A Coruña (0.032), Cáceres (0.035) and Zamora, Pontevedra, Seville and Cádiz (0.036), all below 4%, and located rather in the peninsular East. Table A1 (see Appendix A) shows a more detailed evolution of these figures for each year between 2015 and 2020.

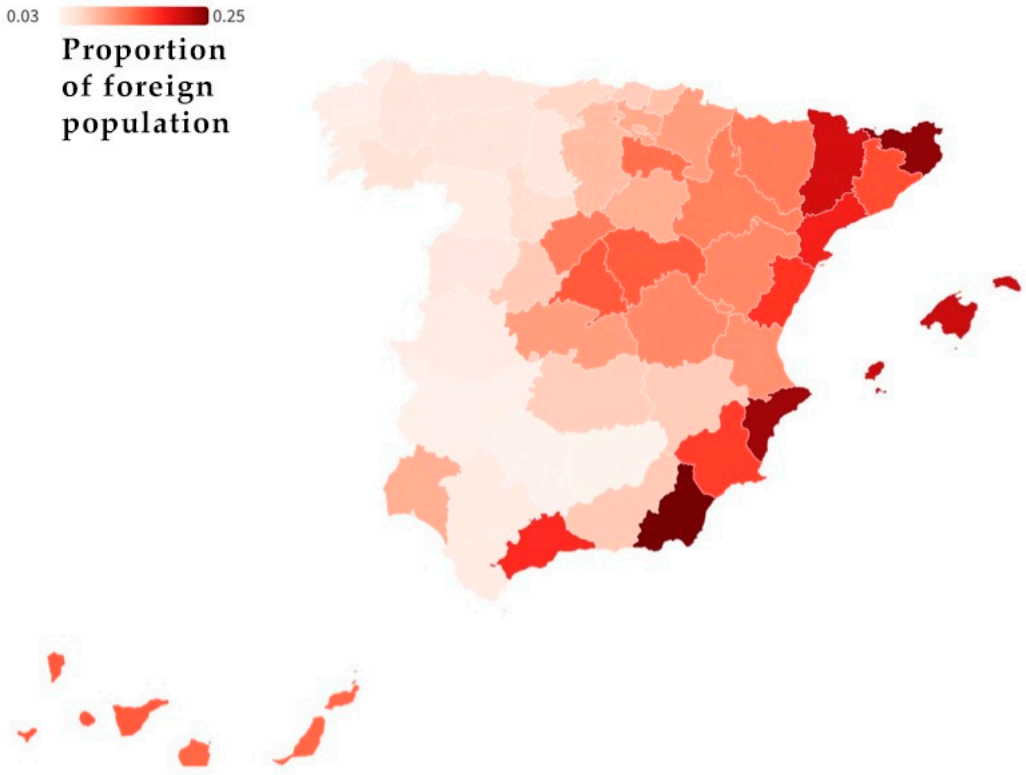

Figure 1. Average proportion of foreign population in Spain (2015–2020). Own elaboration, based on data from the Instituto Nacional de Estadística (2022).

The other element to be considered in this first hypothesis is the citizen's acceptance toward immigration, which is reflected in Figure 2 using data from the CIS survey in 2017. The general attitudes are rather positive, with Asturias, Santa Cruz de Tenerife and Navarra as the provinces with a stronger support for immigration, with a positive value over 70%, whereas the provinces with more negative values are Ciudad Real, Lleida, Ceuta, Murcia and Almería, which do not reach 0.5. Table A2 (see Appendix A) shows all the data more in detail.

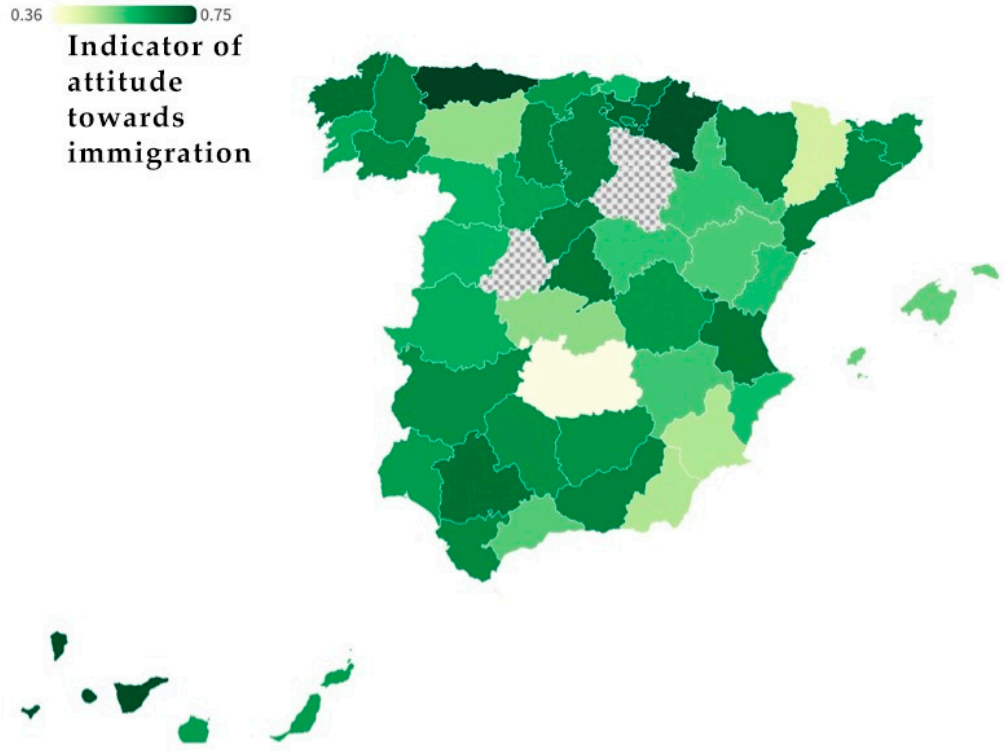

Figure 2. Average acceptance towards immigration in Spain. Own elaboration, using CIS data (2017).

When performing the correlation test between the proportion of foreign population in Spain and the attitude towards immigration, no statistically significant relationship was found (r = 0.067, $p > 0.05$), thus not confirming the validity of H1 (Figure 3).

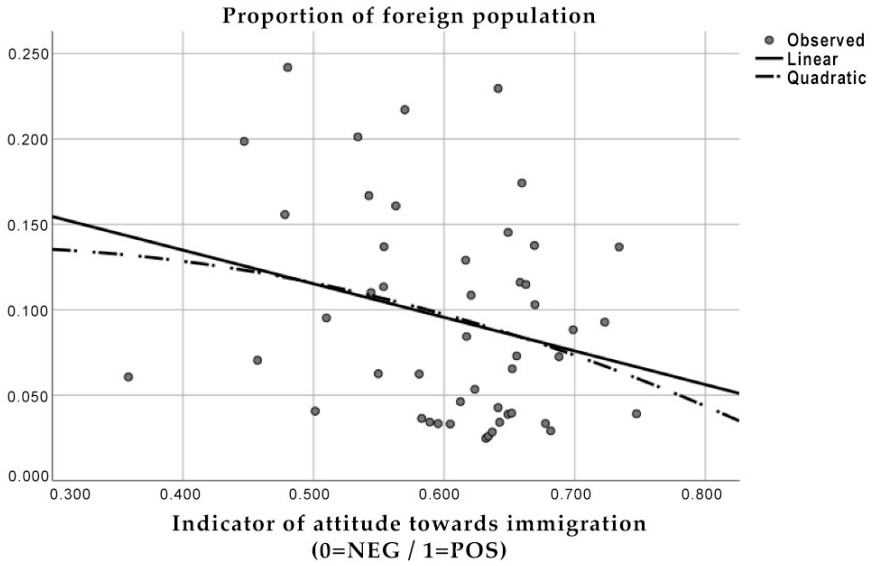

Figure 3. Scatter plot. Correlation between the proportion of foreign population in Spain and the attitude towards immigration. Own elaboration.

We can now move on to H2, in which it was stated that the proportion of immigrant population in Spain at the province level (NUTS 3) is negatively related level of hate speech present on Twitter towards migrants and refugees. In the first place, analyzing a total of 97,710 geolocated Twitter messages in Spain with the hate detector, the average presence of hate speech could be obtained, showing the following average values during each year of the studied period: 2015 = 0.302; 2016 = 0.286; 2017 = 0.270; 2018 = 0.292; 2019 = 0.312; 2020 = 0.308. With this, we can place the total average level of hate against migrants and refugees in tweets (2015–2020) in Spain at 0.29 (SD = 0.03) (see Figure 4 for the province distribution and Table A3 in the Appendix A for more detailed data for each province and year).

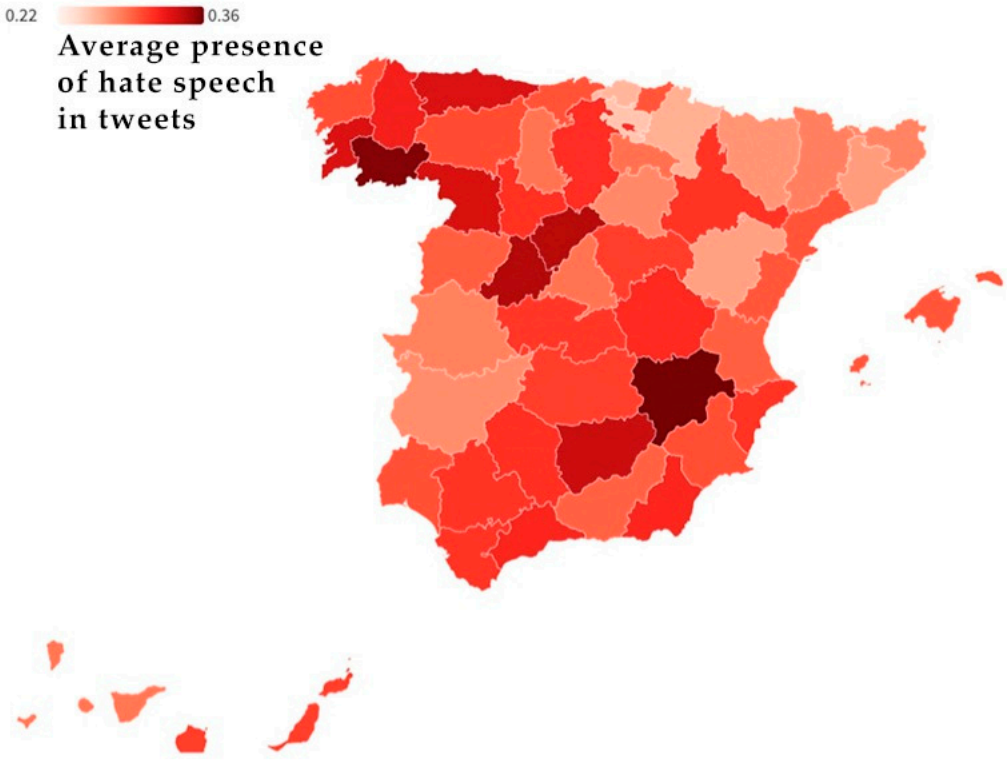

Figure 4. Hate speech total average present on Twitter in Spain. Own elaboration.

The data in Table A3 (see Appendix A) shows some relevant differences between 2015 and 2020, with the central years being the ones with lower levels of hate, and 2015 and 2019 and 2020 with higher levels. At the same time, the highest individual values can be found in 2020 in Guadalajara (M = 0.446), Pontevedra (M = 0.443) and Ourense (M = 0.423). At the same time, also Guadalajara, in 2017, shows the lowest level of all (M = 0.176), followed by Ceuta in 2018 (M = 0.183). Figure 5 shows a visual map with the representation of the average presence of hate speech on Twitter in each Spanish province each year from 2015 to 2020, the most intense color being the one that indicates a higher average of hate speech.

With this, we cannot confirm H2 because the relationship between immigrant population and presence of hate speech variables is not statistically significant ($r = 0.035$, $p > 0.05$). Similar results are obtained in a more detailed study, observing the correlation of the variables year by year, since no significant results are found either: 2015 ($r = 0.015$, $p > 0.05$), 2016 ($r = 0.009$, $p > 0.05$), 2017 ($r = 0.004$, $p > 0.05$), 2018 ($r = 0.005$, $p > 0.05$), 2019 ($r = 0.027$, $p > 0.05$), and 2020 ($r = 0.018$, $p > 0.05$) (Figure 6).

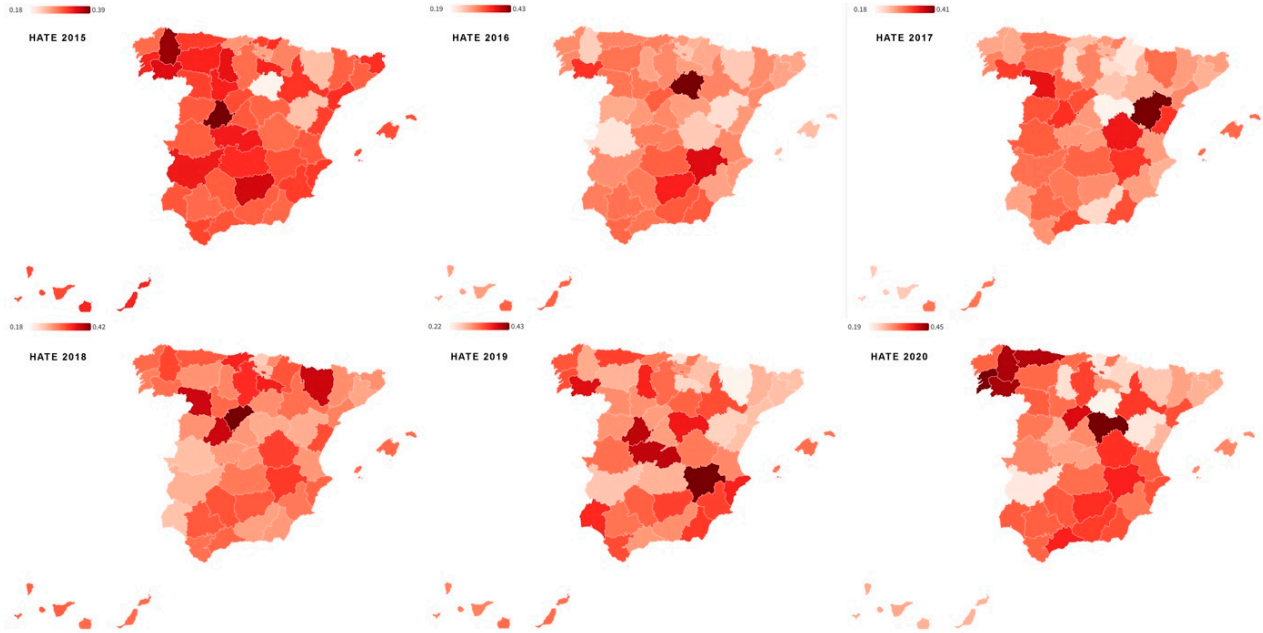

Figure 5. Average presence of hate speech against migrants and refugees on Twitter in Spain by provinces and years. Own elaboration.

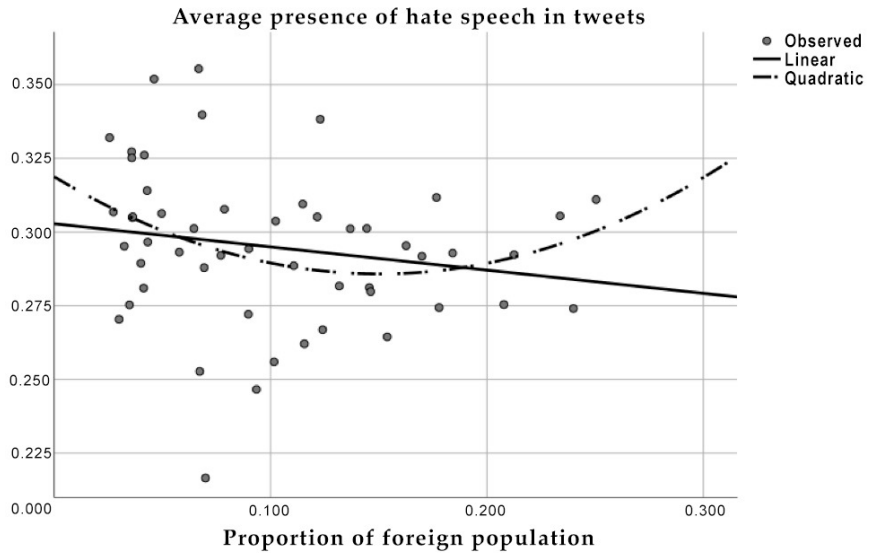

Figure 6. Scatter plot. Correlation between immigrant population and presence of hate speech. Own elaboration.

Finally, we can review the third hypothesis (H3) in which it is stated that the attitude towards immigration and the level of hate speech found on Twitter in this period would be negatively correlated. Using the aforementioned data about both variables, we can see that the results, once again, are not statistically significant ($r = 0.27$, $p > 0.05$), which means that the hypothesis cannot be confirmed (Figure 7).

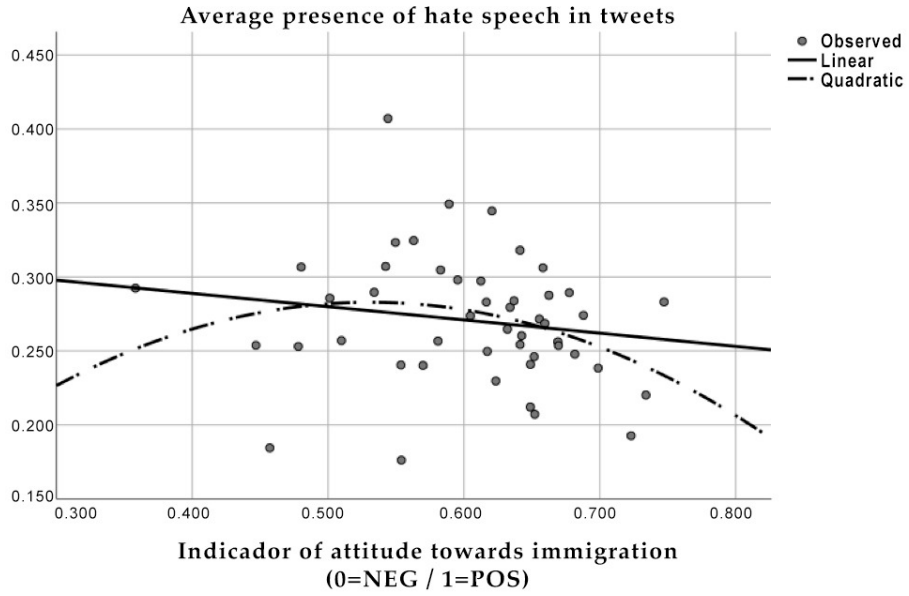

Figure 7. Scatter plot. Correlation between the attitude towards immigration and the level of hate speech found on Twitter. Own elaboration.

4. Discussion

The main observation of the study is the fact that none of the three hypotheses could be confirmed. More explicitly put, it cannot be stated that a greater proportion of immigrants in a province correlates with a greater positive attitude towards this group (H1). It can also not be affirmed that a greater proportion of foreign population in a Spanish province correlates with a lower presence of hate speech on Twitter towards migrants and refugees (H2). Finally, there is no significant negative correlation between the presence of hate speech on Twitter against migrants and refugees and the level of acceptance towards immigration in Spanish provinces (H3).

These observations are not strong enough to refute previous work, such as the one of Arcila-Calderón et al. (2022b), nor the well-established Intergroup and Mediated Intergroup Contact Theory of Allport (1954). They just point out that smaller geographical levels—such as the Spanish province, equivalent to the NUTS 3—might not be the most adequate scenarios for their application, although also this should be further justified.

Some explanations for this result might be precisely related to the size of the population in these entities; some of them with even less than 100,000 inhabitants. First of all, the size of the sample in these provinces when social surveys are conducted is very limited; in this case, the survey conducted by the CIS in 2017 and used as a reference for the measure of the attitudes towards immigrations had no disaggregated data for four of the entities—three provinces and one autonomous city—and in others this data were not big enough to be considered representative. However, no better option with disaggregated and valid data at this NUTS 3 was found, which reinforces the discussion about the limitations of surveys and the need for alternatives that can reach in a more adequate manner smaller levels of analysis.

Accordingly, it should be also noted that a sample of 97,710 tweets, even though relevant for the analysis conducted here, could be increased in further analysis, given that studies using big data approaches are capable of computing much larger volumes of information, thus obtaining also more relevant results. Especially in a case such as the present one, in which the sample is then further divided by year, but also by 52 provinces or autonomous cities, the size in each of the cases rapidly decreases, which makes it necessary for further studies to confirm the observations made here.

Furthermore, it could be also argued that the main differences regarding hate speech and, especially, the attitudes towards immigration, can be found at a national (see, for

instances, the aforementioned Eurobarometers) or, at best, at a regional level (Arcila-Calderón et al. 2022b), but not so much at a provincial level. Thus, even though the hypotheses were not confirmed, our results provide relevant knowledge, as they could indicate that measures to tackle hate speech and attitudes towards immigration might need to have at least regional approaches. However, one of the findings of the studies developed by Arcila-Calderón et al. (2022a, 2022b) is that one of the variables with the greatest predictive weight when developing the models that allowed estimating support for refugees is the country, that is, belonging to a specific country with its peculiar social, political and economic context. That is why the greatest differences were observed directly between the regions of different countries. This could explain why there are no significant differences among the provinces within a country.

It should be also mentioned that the choice of studying migrants and refugees together might limit the validity of the observations, as the attitudes toward both groups are significantly different (Verkuyten et al. 2018), and so is the expression of hate speech and rejection toward both groups on Twitter (Arcila-Calderón et al. 2020a). The choice of studying them together obeys the intention of replicating the approach of Arcila-Calderón et al. (2022b), as well as the intention not to keep reducing sample sizes. Regardless, future studies will need to observe whether these differences also impact the prediction capacities of the model presented here.

Furthermore, the study has exclusively used correlations to test the validity of the hypotheses because they offer a preliminary study that could later lead to more powerful tests to study causality. In this regard, it is important to acknowledge the lack of capacity of stablishing causality interpretations in the study.

Beyond the interpretation regarding the lack of acceptance of our postulated hypotheses, the article has provided relevant knowledge about the conversation about migrants and refugees and about the presence of hate speech against these groups on Twitter in the different Spanish provinces and about its evolution between 2015 and 2020. The first year was the one in which the Refugee Crisis of the Mediterranean had its greatest media presence, and also the moment in which the coverage of immigration experienced a significant change in many European—including Spanish—media (Amores et al. 2019, 2020; Zhang and Hellmueller 2017). The last two years, even including the effect of the pandemic in 2020, which reduced the volume of international human mobility, were the ones with the strongest presence of hate, something that could be associated, among other reasons, to the arrival of Vox to the institutions, introducing in the Spanish public opinion the topic of migration (Castromil et al. 2020; Ferreira 2019), something that until then had not been as relevant and divisive as in other European countries.

Observing the distribution of hate in the different provinces, it is worth mentioning that neither the provinces with a greater number of arrivals—the Canary Islands, the cities of Ceuta and Melilla or the south of Andalusia—nor those with the largest support for anti-immigration parties and policies, show particularly high volumes of hate speech. In general, no relevant trend based on the size of the population, the immigrant proportion in it, the regional division or the party in the government was detected that could explain the different volumes of hate speech. One observation that can be made is that it is usually in less populated ones where the highest, but also the lowest, levels of hatred are found. Although multiple social interpretations could help explaining this, the limited size of the sample in these cases might also play a role, given that a small number of tweets could have a strong influence and, with it, a distorting effect. Further studies with longer periods of analysis and larger sample sizes would be needed to further understand this. Regardless, this study has the particularity of using only geolocated tweets which, on the one hand, significantly reduces the volume of messages to be analyzed (especially in less populated provinces), but at the same time, it ensures their geographical location, which allows this type of analysis to be possible.

Furthermore, it can be considered a limitation of the study the fact that the variables that have a stronger influence on the attitudes towards migration, as well as the level of anti-immigration hatred spread online, have a national dimension. Therefore, it is understood that when descending to these lower and more concrete levels of analysis, it is more difficult to find significant differences. On the other hand, it is true that the main variable (the level of hate speech online against immigration) should not be affected by the level of aggregation, since this level is measured for each one of the tweets. However, it is possible that there are other variables that become more important at these lower and more specific levels of aggregation. Therefore, in subsequent works, the study can be extended to other dimensions that affect the level of hatred and citizen attitudes. The overall analysis might be less powerful due to the low number of cases, compared to the study by Arcila-Calderón et al. (2022b), in which all European regions were considered. However, in this case, the 52 Spanish provinces and 97,719 messages have been considered, which should be sufficient to carry out this type of statistical analysis. Lastly, the nature and scope of this study is exploratory and descriptive, as it is the only approach and methodological strategy that allows analyzing, observing and describing such a large-scale phenomenon. In the future, in order to establish a reliable and confirmatory empirical explanation, it would be necessary to carry out other types of studies, such as those of an experimental nature, to better control other variables and possible confounding factors.

Supplementary Materials: The downloaded data from each tweet can be consulted in https://doi.org/10.6084/m9.figshare.16708942.v1; The keywords used for the download of the tweets at: https://doi.org/10.6084/m9.figshare.16708945.v1.

Author Contributions: Conceptualization, P.S.-H., J.J.A. and D.B.-H.; methodology, P.S.-H. and J.J.A.; software, P.S.-H. and J.J.A.; validation, P.S.-H., J.J.A. and D.B.-H.; formal analysis, P.S.-H., J.J.A. and D.B.-H.; investigation, P.S.-H., J.J.A. and D.B.-H.; resources, P.S.-H., J.J.A. and D.B.-H.; data curation, P.S.-H., J.J.A. and D.B.-H.; writing—original draft preparation, P.S.-H., J.J.A. and D.B.-H.; writing—review and editing, D.B.-H.; visualization, P.S.-H.; supervision, D.B.-H.; project administration, P.S.-H.; funding acquisition, P.S.-H. and D.B.-H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project "Enhanced migration measures from a multidimensional perspective (HumMingBird)" funded by the European Union under the Horizon 2020 Research and Innovation Programme, with reference number 870661. It has also been supported by the project "Preventing Hate Against Refugees and Migrants (PHARM)" funded by the European Union under the Rights, Equality and Citizenship Programme (2014–2020) with reference number 875217 and by the Stop-Hate prototype, funded by the Fundación General de la Universidad de Salamanca as a competitive proof of concept within the TCUE plan, with reference PC-TCUE18-20_016.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to their restricted access only using Twitter's API.

Acknowledgments: The authors would like to thank all the members of the projects PHARM, HumMingBird and StopHate that have made possible the research that is presented here, and in particular, they thank Carlos Arcila Calderón for his leadership.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Percentage of foreign population in Spain [1] by provinces (NUTS 3).

Province	2015	2016	2017	2018	2019	2020	Mean
A Coruña	0.030	0.029	0.029	0.032	0.035	0.040	0.032
Álava	0.087	0.085	0.088	0.094	0.100	0.107	0.094
Albacete	0.068	0.065	0.063	0.065	0.068	0.073	0.067
Alicante	0.245	0.230	0.217	0.225	0.237	0.249	0.234
Almería	0.245	0.245	0.242	0.245	0.256	0.270	0.250
Asturias	0.041	0.040	0.039	0.041	0.042	0.047	0.042
Ávila	0.075	0.070	0.068	0.061	0.064	0.072	0.068
Badajoz	0.031	0.030	0.028	0.029	0.030	0.032	0.030
Baleares	0.211	0.206	0.201	0.206	0.218	0.233	0.213
Barcelona	0.141	0.141	0.145	0.152	0.164	0.180	0.154
Burgos	0.079	0.077	0.073	0.075	0.080	0.087	0.079
Cáceres	0.036	0.035	0.033	0.034	0.035	0.036	0.035
Cádiz	0.035	0.034	0.034	0.035	0.038	0.041	0.036
Cantabria	0.057	0.055	0.053	0.056	0.060	0.066	0.058
Castellón	0.176	0.169	0.161	0.164	0.170	0.180	0.170
Ceuta	0.064	0.068	0.070	0.073	0.072	0.072	0.070
Ciudad Real	0.072	0.065	0.061	0.061	0.062	0.067	0.065
Córdoba	0.028	0.026	0.026	0.027	0.028	0.030	0.027
Cuenca	0.114	0.111	0.109	0.111	0.117	0.128	0.115
Girona	0.236	0.232	0.230	0.235	0.247	0.259	0.240
Granada	0.069	0.067	0.065	0.067	0.071	0.076	0.069
Guadalajara	0.145	0.140	0.137	0.140	0.147	0.158	0.144
Guipúzcoa	0.069	0.071	0.072	0.076	0.083	0.090	0.077
Huelva	0.088	0.086	0.084	0.088	0.094	0.099	0.090
Huesca	0.120	0.118	0.115	0.119	0.129	0.144	0.124
Jaén	0.027	0.026	0.025	0.024	0.025	0.027	0.026
La Rioja	0.132	0.127	0.125	0.128	0.134	0.145	0.132
Las Palmas	0.136	0.131	0.129	0.135	0.141	0.149	0.137
León	0.045	0.042	0.041	0.042	0.043	0.047	0.043
Lleida	0.206	0.204	0.199	0.203	0.212	0.224	0.208
Lugo	0.039	0.039	0.039	0.042	0.046	0.052	0.043
Madrid	0.143	0.138	0.138	0.142	0.151	0.162	0.146
Málaga	0.180	0.173	0.167	0.170	0.180	0.191	0.177
Melilla	0.174	0.179	0.179	0.181	0.177	0.177	0.178
Murcia	0.165	0.159	0.156	0.159	0.165	0.172	0.163
Navarra	0.096	0.093	0.093	0.100	0.109	0.120	0.102
Ourense	0.044	0.043	0.043	0.045	0.048	0.054	0.046
Palencia	0.040	0.039	0.039	0.041	0.043	0.047	0.041
Pontevedra	0.035	0.033	0.033	0.035	0.038	0.042	0.036
Salamanca	0.039	0.038	0.036	0.039	0.042	0.047	0.040
Santa Cruz de Tenerife	0.139	0.139	0.137	0.145	0.155	0.163	0.146
Segovia	0.126	0.121	0.116	0.118	0.124	0.133	0.123
Sevilla	0.037	0.035	0.033	0.035	0.037	0.041	0.036
Soria	0.089	0.087	0.082	0.086	0.092	0.103	0.090
Tarragona	0.184	0.178	0.174	0.179	0.189	0.201	0.184
Teruel	0.119	0.113	0.110	0.112	0.116	0.124	0.116
Toledo	0.104	0.099	0.095	0.097	0.104	0.115	0.102
Valencia	0.108	0.107	0.103	0.107	0.114	0.125	0.111
Valladolid	0.051	0.049	0.046	0.047	0.050	0.055	0.050
Vizcaya	0.060	0.061	0.062	0.066	0.072	0.081	0.067
Zamora	0.038	0.037	0.034	0.034	0.035	0.037	0.036
Zaragoza	0.119	0.112	0.113	0.117	0.129	0.140	0.122

[1] Continuous variable from 0 to 1, with 0 meaning the absence of foreigners in the province and 1 meaning that the entire population has foreign origins. Source: own elaboration based on data from the Instituto Nacional de Estadística (2022).

Table A2. Average acceptance towards immigration in Spain [1] by provinces (NUTS 3).

Province	Attitude
Ávila	-
La Rioja	-
Melilla	-
Soria	-
Asturias	0.747
Santa Cruz de Tenerife	0.734
Navarra	0.723
Álava	0.699
Guipúzcoa	0.688
A Coruña	0.682
Sevilla	0.678
Valencia	0.669
Madrid	0.669
Huesca	0.663
Tarragona	0.660
Segovia	0.658
Burgos	0.655
Granada	0.652
Lugo	0.652
Barcelona	0.649
Palencia	0.649
Cádiz	0.642
Girona	0.641
Ourense	0.641
Badajoz	0.637
Córdoba	0.634
Jaén	0.632
Cantabria	0.623
Cuenca	0.621
Huelva	0.617
Las Palmas	0.616
Valladolid	0.612
Pontevedra	0.605
Cáceres	0.595
Zamora	0.589
Salamanca	0.583
Vizcaya	0.581
Alicante	0.570
Castellón	0.563
Guadalajara	0.554
Zaragoza	0.554
Albacete	0.550
Teruel	0.544
Málaga	0.542
Baleares	0.534
Toledo	0.510
León	0.501
Almería	0.480
Murcia	0.478
Ceuta	0.457
Lleida	0.447
Ciudad Real	0.358

[1] Continuous variable from 0 to 1, with 0 being the most negative attitude towards migration and 1 being the most positive attitude or greater support for migrants. Source: own elaboration using CIS data (2017).

Table A3. Hate speech average present on Twitter in Spain [1] by provinces (NUTS 3).

Province	2015	2016	2017	2018	2019	2020	Total Mean
A Coruña	0.285	0.284	0.248	0.298	0.333	0.323	0.295
Álava	0.251	0.242	0.238	0.236	0.298	0.215	0.247
Albacete	0.301	0.379	0.323	0.330	0.429	0.370	0.355
Alicante	0.314	0.275	0.240	0.302	0.366	0.335	0.305
Almería	0.287	0.319	0.307	0.265	0.351	0.338	0.311
Asturias	0.313	0.290	0.283	0.309	0.345	0.416	0.326
Ávila	0.392	0.275	0.326	0.380	0.402	0.264	0.340
Badajoz	0.335	0.282	0.284	0.249	0.263	0.210	0.270
Baleares	0.302	0.248	0.290	0.299	0.320	0.296	0.292
Barcelona	0.291	0.267	0.241	0.257	0.268	0.264	0.264
Burgos	0.284	0.287	0.272	0.338	0.322	0.345	0.308
Cáceres	0.293	0.211	0.298	0.237	0.318	0.295	0.275
Cádiz	0.306	0.310	0.260	0.293	0.340	0.322	0.305
Cantabria	0.260	0.294	0.230	0.345	0.314	0.316	0.293
Castellón	0.310	0.275	0.325	0.318	0.266	0.258	0.292
Ceuta	0.229	0.191	0.185	0.183	0.292	0.220	0.217
Ciudad Real	0.322	0.315	0.292	0.287	0.276	0.315	0.301
Córdoba	0.294	0.296	0.280	0.314	0.328	0.330	0.307
Cuenca	0.284	0.236	0.345	0.324	0.312	0.357	0.310
Girona	0.320	0.266	0.254	0.272	0.266	0.266	0.274
Granada	0.292	0.320	0.207	0.260	0.300	0.349	0.288
Guadalajara	0.289	0.272	0.176	0.253	0.372	0.446	0.301
Guipúzcoa	0.322	0.256	0.274	0.275	0.315	0.311	0.292
Huelva	0.297	0.295	0.250	0.235	0.362	0.328	0.294
Huesca	0.232	0.237	0.288	0.380	0.219	0.245	0.267
Jaen	0.355	0.359	0.265	0.309	0.345	0.359	0.332
La Rioja	0.327	0.271	0.206	0.350	0.251	0.287	0.282
Las Palmas	0.326	0.315	0.283	0.298	0.321	0.263	0.301
León	0.328	0.302	0.286	0.270	0.276	0.318	0.297
Lleida	0.291	0.284	0.254	0.273	0.275	0.275	0.275
Lugo	0.379	0.234	0.246	0.322	0.282	0.422	0.314
Madrid	0.295	0.286	0.256	0.267	0.289	0.293	0.281
Málaga	0.299	0.301	0.307	0.302	0.293	0.370	0.312
Melilla	0.287	0.316	0.269	0.255	0.288	0.233	0.274
Murcia	0.312	0.267	0.253	0.288	0.345	0.307	0.295
Navarra	0.270	0.263	0.193	0.312	0.267	0.231	0.256
Ourense	0.346	0.347	0.318	0.295	0.383	0.423	0.352
Palencia	0.341	0.267	0.212	0.280	0.365	0.221	0.281
Pontevedra	0.327	0.286	0.274	0.283	0.339	0.443	0.325
Salamanca	0.294	0.263	0.305	0.271	0.297	0.307	0.289
Santa Cruz de Tenerife	0.304	0.271	0.220	0.306	0.308	0.270	0.280
Segovia	0.303	0.317	0.306	0.419	0.290	0.395	0.338
Sevilla	0.286	0.303	0.289	0.310	0.317	0.325	0.305
Soria	0.184	0.431	0.225	0.273	0.326	0.194	0.272
Tarragona	0.320	0.285	0.269	0.283	0.266	0.335	0.293
Teruel	0.229	0.218	0.407	0.251	0.255	0.213	0.262
Toledo	0.333	0.293	0.257	0.265	0.400	0.275	0.304
Valencia	0.296	0.286	0.254	0.266	0.303	0.327	0.289
Valladolid	0.330	0.300	0.297	0.299	0.297	0.315	0.306
Vizcaya	0.309	0.275	0.257	0.228	0.237	0.211	0.253
Zamora	0.309	0.301	0.349	0.379	0.301	0.325	0.327
Zaragoza	0.317	0.290	0.241	0.294	0.339	0.351	0.305

[1] Continuous variable from 0 to 1, with 0 being the absence of anti-immigration hate and 1 being the presence of explicit hate. Source: own elaboration.

Note

[1] The graphical interface is available online at http://pharm-interface.usal.es (accessed on 8 November 2022).

References

Abrams, Dominic, and Michael A. Hogg. 2017. Twenty years of group processes and intergroup relations research: A review of past progress and future prospects. *Group Processes & Intergroup Relations* 20: 561–69. [CrossRef]

Allport, Gordon W. 1954. *The Nature of Prejudice*. Reading: Addison-Wesley.

Amores, Javier J., Carlos Arcila-Calderón, and David Blanco-Herrero. 2020. Evolution of negative visual frames of immigrants and refugees in the main media of Southern Europe. *Profesional de la Información* 29: e290624. [CrossRef]

Amores, Javier J., Carlos Arcila-Calderón, and Mikolaj Stanek. 2019. Visual frames of migrants and refugees in the main Western European media. *Economics and Sociology* 12: 147–61. [CrossRef]

Arcila-Calderón, Carlos, David Blanco-Herrero, and María-Belén Valdez-Apolo. 2020a. Rechazo y discurso de odio en Twitter: Análisis de contenido de los tuits sobre migrantes y refugiados en español. *Revista Española de Investigaciones Sociológicas* 172: 21–40. [CrossRef]

Arcila-Calderón, Carlos, Gonzalo de la Vega, and David Blanco-Herrero. 2020b. Topic Modeling and Characterization of Hate Speech against Immigrants on Twitter around the Emergence of a Far-Right Party in Spain. *Social Sciences* 9: 188. [CrossRef]

Arcila-Calderón, Carlos, Javier J. Amores, and M. Mikolaj Stanek. 2022a. Predicting integration of refugees: Using machine learning and synthetic populations to predict social acceptance of asylum seekers in European regions. In *Data Science for Migration and Mobility*. Edited by lbert Ali Salah, Emre Eren Korkmaz and Tuba Bircan. Oxford: Oxford University Press.

Arcila-Calderón, Carlos, Patricia Sánchez-Holgado, Cristina Quintana-Moreno, Javier J. Amores, and David Blanco-Herrero. 2022b. Discurso de odio y aceptación social hacia migrantes en Europa: Análisis de tuits con geolocalización. *Comunicar: Revista Científica de Comunicación y Educación* 30: 21–35. [CrossRef]

Bartlett, Jamie, Jeremy Reffin, Noelle Rumbale, and Sarah Williamson. 2014. *Anti-Social Media*. London: Demos.

Castromil, Antón R., Raquel Rodríguez-Díaz, and Paula Garrigós. 2020. La agenda política en las elecciones de abril de 2019 en España: Programas electorales, visibilidad en Twitter y debates electorales. *El Profesional de la Información* 29: e290217. [CrossRef]

Cea d'Ancona, María Ángeles. 2009. La compleja detección del racismo y la xenofobia a través de encuesta. Un paso adelante en su medición. *Revista Española de Investigaciones Sociológicas (REIS)* 125: 13–45.

Chaudhry, Irfan. 2015. Hashtagging Hate: Using Twitter to Track Racism Online. *First Monday* 20. [CrossRef]

Council of Europe. 1997. Recommendation No. R (97) 20 of the COMMITTEE of Ministers to Member States on "Hate Speech". Available online: https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680505d5b (accessed on 8 November 2022).

Council of the European Union. 2008. Council Framework Decision 2008/913/JHA of 28 November 2008 on Combating Certain Forms and Expressions of Racism and Xenophobia by Means of Criminal Law. Available online: https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:32008F0913 (accessed on 8 November 2022).

Davidson, Thomas, Dana Warmsley, Michael Macy, and Ingmar Weber. 2017. Automated Hate Speech Detection and the Problem of Offensive Language. Paper presented at Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017), Montreal, QC, Canada, May 15–18; Available online: http://sdl.soc.cornell.edu/img/publication_pdf/hatespeechdetection.pdf (accessed on 8 November 2022).

Esses, Victoria M., John F. Dovidio, Antoinette H. Semenya, and Lynne M. Jackson. 2005. Attitudes Towards Immigrants and Immigration: The Role of National and International Identity. In *The Social Psychology of Inclusion and Exclusion*. Edited by Dominic Abrams, Michael A. Hogg and José M. Marques. New York: Psychology Press, pp. 317–37.

European Commission. 2015. Special Eurobarometer 437. Discrimination in the European Union 2015. Available online: https://europa.eu/eurobarometer/surveys/detail/2077 (accessed on 8 November 2022).

European Commission. 2018. Special Eurobarometer 469. Integration of Immigrants in the European Union. Available online: https://europa.eu/eurobarometer/surveys/detail/2169 (accessed on 8 November 2022).

European Commission. 2019. Special Eurobarometer 493. Discrimination in the European Union. Available online: https://europa.eu/eurobarometer/surveys/detail/2251 (accessed on 8 November 2022).

European Commission. 2022. Special Eurobarometer 519. Integration of Immigrants in the European Union. Available online: https://europa.eu/eurobarometer/surveys/detail/2276 (accessed on 8 November 2022).

Ferreira, Carles. 2019. Vox como representante de la derecha radical en España: Un estudio sobre su ideología. *Revista Española de Ciencia Política* 51: 73–98. [CrossRef]

Instituto Nacional de Estadística. 2022. Población extranjera por Nacionalidad, comunidades, Sexo y Año. Available online: https://www.ine.es/jaxi/Tabla.htm?path=/t20/e245/p08/l0/&file=02005.px&L=0 (accessed on 8 November 2022).

Kreis, Ramona. 2017. #refugeesnotwelcome: Anti-refugee Discourse on Twitter. *Discourse & Communication* 11: 498–514. [CrossRef]

Miró Llinares, Fernando. 2016. Taxonomía de la comunicación violenta y el discurso del odio en internet. *IDP: Revista De Internet, Derecho Y Política* 22: 82–107. [CrossRef]

Müller, Karsten, and Carlo Schwarz. 2020. Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association* 19: 2131–67. [CrossRef]

Organization for Security and Co-operation in Europe. n.d. OSCE-ODIHR Hate Crime Reporting. Available online: https://hatecrime.osce.org (accessed on 8 November 2022).

Tajfel, Henri. 1978. *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations*. New York: Academic Press.

Turnbull-Dugarte, Stuart J. 2019. Explaining the end of Spanish exceptionalism and electoral support for Vox. *Research & Politics* 6: 1–8. [CrossRef]

Verkuyten, Maykel, Kieran Mepham, and Mathijs Kros. 2018. Public Attitudes towards Support for Migrants: The Importance of Perceived Voluntary and Involuntary Migration. *Ethnic and Racial Studies* 41: 901–18. [CrossRef]

Vrysis, Lazaros, Nikolaos Vryzas, Rigas Kotsakis, Theodora Saridou, Maria Matsiola, Andreas Veglis, Carlos Arcila-Calderón, and Charalampos Dimoulas. 2021. A Web interface for analyzing hate speech. *Future Internet* 13: 80. [CrossRef]

Zhang, Xu, and Lea Hellmueller. 2017. Visual framing of the European refugee crisis in Der Spiegel and CNN International: Global journalism in news photographs. *The International Communication Gazette* 79: 483–510. [CrossRef]