

Politicization of Immigration and Language Use in Political Elites: A Study of Spanish Parliamentary Speeches

Journal of Language and Social Psychology

1–31

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0261927X231175856

journals.sagepub.com/home/jls

Berta Chulvi^{1,2} , Mariangeles Molpeceres², María F. Rodrigo³, Alejandro H. Toselli¹, and Paolo Rosso¹

Abstract

This study uses natural language processing (NLP) tools to analyze how politicians recreate the stereotype of immigrants in the Spanish Parliament. An interdisciplinary approach from computational linguistics and social psychology has been used to construct a variety of indices about content and linguistic styles. The analysis of 2,516 parliamentary interventions about immigration delivered between 1996 and 2016 by representatives of the two political parties that alternated in power during that period (conservative Popular Party and Spanish Socialist Party) shows that both the rhetorical strategy to present immigrants as “victims” or as a “threat” and the language style that politicians use reveal an interaction between the ideology of the party and the party’s political position in government or in the opposition. Results also suggest some changes over time in the polarization and politicization of the debate about immigration.

Keywords

immigration, stereotyping, prejudice, parliamentary discourse, computational linguistics

This paper approaches Spanish parliamentary discourse about immigration, focusing on both stereotype content and subtler language usage indices as related to party ideology and political position in government or in the opposition. Drawing upon a corpus

¹Universitat Politècnica de València, Valencia, Spain

²Social Psychology Department, Universitat de València, Valencia, Spain

³Methodology of the Behavioural Sciences Department, Universitat de València, Valencia, Spain

Corresponding Author:

Berta Chulvi, Universitat Politècnica de València, Valencia, Spain.

Email: berta.chulvi@upv.es

of parliamentary speeches in Spanish Congress between 1996 and 2016, our study explores how both party ideology and political position influence the use of different frames to stereotype immigrant persons, as well other subtler linguistic indices that have previously been found related to outgroup negativity. While political discourse on immigration has been a hot issue in European political science research in recent years, countries in Southern Europe, which have a more recent history of immigration, have been largely overlooked in this research. However, their consideration might be enlightening as to the role that mainstream parties have played in the politicization of immigration in Europe.

Ideology and Party Competition: Politicization of Immigration in Europe

Over the past decades, immigration has become one of the most salient and controversial issues in party competition in Western Europe. There is plenty of evidence linking immigration attitudes to right-wing or left-wing ideological orientations when examining attitudes of both political elites (Alonso & Fonseca, 2011; Kocijan & Kuček, 2022) and the general public (Brooks et al., 2016). There is also some empirical evidence that migration inflows influence both political parties' and general public's attitudes to immigration (Consterdine & Kintz, 2018; Markaki & Longhi, 2013). However, neither parties' ideology nor inflow data are enough to account for the more or less restrictive stance that European political parties adopt toward immigration (Haas & Natter, 2015; Odmalm, 2012).

This is why there has been abundant discussion in recent years about immigration '*politicization*' in Europe. This literature suggests that, besides ideological preferences and whatever constraints migration inflows may impose on elites' discourses, political parties adopt strategic stances on immigration issues depending on the context of party competition (for a review see Van der Brug et al., 2015). An issue such as immigration is said to be politicized when (i) it suddenly acquires high visibility or *salience* in the political arena and (ii) there is an increasing *polarization* or disagreement in the position that political actors adopt on the topic (De Wilde, 2011). Van der Brug et al. (2015) suggest that (a) ideologically loaded topics, such as immigration, are more likely to be politicized than topics cutting across existing coalitions, and that (b) challengers to parties in power (i.e., opposition or new parties) have more incentives to increase the salience of such divisive issues. Therefore, it would be expected that parties' handling of controversial issues such as immigration would be determined by both their ideological orientation and their position in the political field at a given moment (Odmalm, 2018; Odmalm & Bale, 2015).

Recent interest in the political elites' discourse about immigration has been mainly driven by the emergence of radical right parties regarded as responsible for the politicization of the topic (Hooghe & Marks, 2009; Kriesi et al., 2008; Mudde, 2007). However, some research suggests that the role of mainstream parties may have been largely underestimated or even ignored, while the mobilizing force of radical right

challenger parties may have been overestimated (Abdou et al., 2022; Grande et al., 2019; Odmalm & Bale, 2015). This relative neglect of the role of mainstream parties should be amended, particularly when there are increasing indications that some of the contemporary restrictive turns on immigration may result from parties' competition with their *mainstream* challengers, especially in the case of mainstream right-wing or center-right parties (Odmalm & Bale, 2015; Odmalm, 2018).

Moreover, literature examining determinants of mainstream parties' politics of migration tends to picture political parties as passive agents that primarily react to public opinion, comparatively disregarding parties' agency and thus both (i) mainstream parties' own pro-active reasons for moving in a more open or stricter direction should they identify some sort of potential electoral gain and (ii) the influence that political elites' discourse may have in shaping public opinion itself (Odmalm, 2018). In other words, there are some indications that politicization of immigration in Europe in the twenty-first century may well have been a top-down process, in which mainstream and government parties have played an especially important role (Bale, 2008; Green-Pedersen & Krogstrup, 2008; Van der Brug et al., 2015), with changes in political rhetoric that took place well before recent refugee crises (Abdou et al., 2022). However, little attention has been paid to the electoral incentives or political position of mainstream parties as incumbent or challenger in examining their role in politicization of migration.

Politicization of Immigration in Spain

Spain is no exception to this trend of politicization of immigration in Western Europe, even if the timing and development of such a process show some particularities of the Spanish case that deserve attention. To start with, unlike other European countries such as the U.K., France or the Netherlands, immigration in Spain is a relatively recent phenomenon. The number of foreign individuals in the country increased from 637,085 to 4,618,581 from 1996 to 2016 (see Table A1, in Appendix 1, in Supplemental Material). Foreign-born population amounted to 13.15% in 2016, an impressive change from the less than 4% recorded in 2000. The rate of increase very rapidly accelerated after 2000, multiplying by five and averaging around half a million individuals per year between 2000 and 2008. After 2008, that rate of increase stabilized and then the trend slowly started to reverse (see Table A1, in Appendix 1, in Supplemental Material). This level of growth in such a short period of time is unparalleled in other European countries.

These particularities, together with the also comparatively late emergence of a radical right party as a nation-wide political contender (Dennison & Mendes, 2019; Turnbull-Dugarte, 2019) and a relatively lenient attitude of Spaniards towards immigrants (Mendez et al., 2014), have led some researchers to suggest that Spain might be an exception when it comes to politicization of immigration compared to other European countries (Alonso & Rovira, 2015; Gonzalez-Enriquez, 2017; Rinken, 2015). However, contrary to the illusion of Spanish exceptionalism, recent research (Gattinara & Morales, 2017; Morales et al., 2015; Pardos-Prado, 2015; Ros &

Morales, 2015) suggests that Spanish mainstream parties had started to incorporate immigration into their patterns of electoral competition before the rise of a radical right party, and that Spanish party politics of migration would be quite similar to that found elsewhere in Western Europe (Odmalm, 2018).

In fact, Morales et al. (2015), in a paper examining manifestos and media claims of Spain's two main national parties, conservative PP and moderate-left Socialist Party¹, from 1993 to 2011, find a correlation between salience of the issue and polarization of both parties' stances on immigration, a result that points to the very definition of politicization given in earlier pages (van der Brug et al., 2015). In their own words,

"[...] we found [...] greater attention being paid to immigration when parties were more polarized around the issue, or higher polarization whenever immigration was prominent. It is difficult to say whether polarization drives salience or the other way around. However, a careful reading of the newspaper reports [...] suggests that the salience of the issue forces parties to position themselves on the different aspects related to immigration in order to attend to the news cycle. This, in turn, forces them to become less ambiguous about their positions, thus fostering polarization" (Morales et al., 2015: 25).

Despite the impressive speed of migration inflows from 2000 to 2008, immigration was not a particularly salient concern throughout the whole period, as plotted in Figure 1. Rather occasional peaks of salience of immigration in Spanish public debate concur with particular events that were prominent in the media, such as two

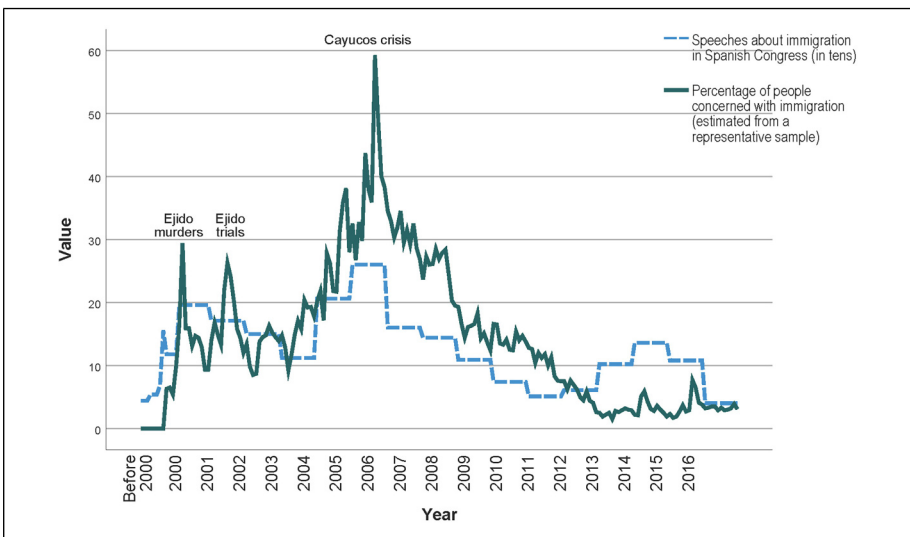


Figure 1. Percentage of people who regard immigration as one of Spain's three main problems (solid line) and number (in tens) of yearly speeches about immigration in Spanish Congress (dashed line). Sources: CIS Barometer for the former and ParlSpeech for the latter. Ç.

murders by Moroccan immigrants at the end of 2000 in Southern Spain that went on trial in 2002. However, during the summer of 2006, a huge humanitarian crisis was triggered by the arrival of more than 18,000 African immigrants in canoes or 'cayucos' to the Canary Islands in less than a month, a figure that surpassed the number of immigrants that had arrived to the archipelago during the full previous year. Such a pressing crisis, that resulted in the whole EU reconsidering its immigration policy, triggered a huge increase of immigration salience in public concern that was mirrored by a parallel increment in Spanish parliamentary debate, as reflected in Figure 1.

Our paper aims to examine how mainstream political elites talked about immigrants in Spain before any radical right party emerged as a national contender². It should be taken into account that some data on public and political concern about the topic suggest that 2006 might have marked a turning point in Spanish political debate on immigration, thrusting the issue into the public limelight as it had never been before since migration inflows started. Exploring whether such sudden salience of immigration in parliamentary debate concurs with a strategic shift of the (then challenger) Popular Party might illuminate factors that shape politicization processes.

Language Use by Politicians and Intergroup Dynamics

Parliamentary discourse, besides its function of legislation and policy-making, also has a function of shaping the attitudes and opinions about immigration of the general public and influencing electoral outcomes through the media coverage that political discourse receives (e.g., Gaucher et al., 2018; Heizmann & Huth, 2021). Our focus here will be on examining the representation of immigrants in parliamentary discourse by means of (i) negative stereotyping of the outgroup in parliamentary debate, and (ii) the use and characterization of ingroup-outgroup categories in it.

Despite their potential importance for understanding the interplay between ideological discourse and party competition, categorical schemata and intergroup representations conveyed by political elites have typically been difficult to study because such a target is difficult to reach through traditional methods of researching stereotypes and intergroup representations. However, such elements may also be studied unobtrusively through the use of text analysis, especially when combined with automated methods of textual analysis (Charlesworth & Banaji, 2022; Jost & Sterling, 2020). Therefore, social psychology and political science have featured in recent years increasing literature on language usage by politicians attending to the new possibilities that computational linguistics offers (Houck et al., 2022; Jordan, 2022; Sterling, Jost & Bonneau, 2020).

Stereotyping and Framing Immigrants

A long tradition of research in social psychology defines stereotypes as a set of widespread beliefs associated with a group category (Dovidio et al., 2010). The common assumption is that these sets of beliefs activate the operation of prejudices and justify them because they build on socially shared cognitive schemata about target

groups. Stereotypes may be transmitted and perpetuated explicitly through language, for instance, by associating group labels with corresponding traits and behaviors (Keblusek et al. 2017). And indeed, mainstream research on stereotypes has typically focused on the examination of the content of these beliefs, as assessed by a list of attributed features. For example, the Stereotype Content Model (SCM) developed by Fiske and colleagues (2007) argues that humans use two universal dimensions of social cognition—warmth and competence—to judge individuals and groups, and it examines stereotype content and valence through the attribution of both features to a variety of target groups (see Lee and Fiske, 2006 for an application of the SCM to immigrants), resulting in different forms of ambivalent prejudice such as menacing or paternalistic prejudice. Some research relying on more inductive methods to explore stereotype content has qualified and complemented SCM, highlighting the significance that perceived vulnerability or potential disruption have in the representation of immigrants (Savaş et al., 2021; Wei et al., 2019).

However, an explicit judgment of competence or niceness of immigrant people is unlikely to be found in the discourse of political elites in Western democracies, and even less so in parliamentary speeches, where talk of ethnic affairs is usually highly self-controlled. In regimes governed by self-declared egalitarian principles, politicians are acutely aware that any express negative judgment or any attribution of a negative trait may lead to angry reactions from minority groups as well as from rival political groups—they may be accused of bias, xenophobia, or even racism as soon as they derogate immigrants or minorities (Van Dijk, 1999). Therefore, subtler ways to detect immigrant stereotyping are required.

This paper uses the concept of *frame* (Goffman, 1974; Lakoff, 2004) to study how stereotypes emerge in political discourse. Framing, or placing a group in a certain narrative scenario, can convey a prejudiced image without explicitly describing the group as such. This rhetorical strategy sidesteps potentially racist or xenophobic characterizations by instead using actions or facts associated with the group (e.g., “they jump the fence,” “they come in droves”) to activate a negative cognitive frame, consolidating a certain image of the group.

Some research in sociology (e.g., Greussing & Boomgaarden, 2017; Kellstedt, 2003), media studies (Benson, 2013) and computational linguistics (Mendelsohn et al., 2021) has applied frame theory to the study of immigration, pointing to a dual tension between humanitarian and security concerns in the way immigrants are framed. However, those pieces of research have mostly focused on mass media or social media data, thus neglecting what Helbling (2014) identifies as the first phase of the framing process—i.e., the construction of information by political elites. Moreover, such research usually draws upon different manually annotated corpora, making it difficult to compare results when annotation guidelines differ or tagging categories do not fully overlap.

To try and advance research in the intersection of social psychology and NLP on immigration stereotyping, Sánchez-Junquera et al. (2021) have recently offered and validated a domain-specific standard to discriminate between “victim-framed” and “threat-framed” stereotypes of immigrants. Drawing upon a frame-theory inspired

taxonomy focusing on narrative scenarios rather than attributes assigned to a group, Sánchez-Junquera et al. (2021) developed an NLP research to discriminate between texts that stereotype immigrants as a *threat* to the host society or as *victims* of xenophobia or other kinds of suffering. In so doing, their intention was to offer a validated standard upon which to develop further research on immigrant's stereotyping that could be used by different researchers and would allow to obtain comparable results.

Other Linguistic Indicators of Intergroup Polarization

While the analyses of politicization of immigration in political discourse and the media have typically focused on content or frames used in such discourse, nowadays there is also clear and increasing evidence of how small differences in language use can affect the perception of people and groups by focusing (or diverting) attention and guiding social inferences in certain directions (Karasawa & Maass, 2008). Beyond semantic analysis, different levels of linguistic abstraction or subtle grammar choices can influence implicit causality or trait inferences about ingroups and outgroups, affecting intergroup evaluation and thus levels of intergroup polarization or conflict (Anolli et al. 2006; Graf et al., 2013; Formanowicz, 2020; Rubini et al., 2014).

One of the most commonly used tools to capture this subtler layer of language in political elites has been the "Linguistic Inquiry and Word Count" (LIWC) (Pennebaker et al., 2007), a dictionary that classifies words in psycholinguistics categories. Many researchers have used it to examine the effect of political ideology on contrasting semantic choices (e.g., Jost & Sterling, 2020; Körner et al., 2022) or differences in the use of function words associated with reasoning styles (e.g., Cichocka et al., 2016).

It is surprising, though, that LIWC has comparatively been less used to examine language choices as related to escalation of political conflict, given that it is well established that function words, and especially grammatical person use, can be illustrative to examine intergroup dynamics (Pennebaker, 2011). In particular, "you" language has been associated in applied communication research with blaming, attacking, aversion, and antagonism (Van Swol & Carlson, 2015), as well as with hostility and willingness to confront in the context of personal relationships (Tausczik & Pennebaker, 2010). In fact, second-person pronouns have been found to be used in the service of accusation so often as to say that one "can think of using 'you' words as the equivalent of pointing your finger at the other person while talking" (Pennebaker, 2011, p. 175). These results in different kinds of interaction suggest that the use of second-person pronouns might also be sensitive to escalation of political antagonism.

Also particularly interesting for our objectives in this paper is research focused on the relation between use of language as captured by the LIWC and outgroup stigmatization, where a consistent relation has been found between the use of pronouns and outgroup negativity. There is a higher frequency of third-person plural pronouns in language used by a variety of extremist groups of different kinds, such as American nazis, animal-rights groups (Seyle & Pennebaker, 2004), violent lone offenders (Kaati et al., 2016a), terrorists (Pennebaker & Chung, 2008), or their supporters (Torregrosa et al., 2020).

Cohen and her colleagues have repeatedly found that the use of third person plurals is indicative of a preoccupation with an opposing group (Kaati et al., 2016a; Kaati et al., 2016b; Shrestha et al., 2020). However, when discussing immigration, a higher frequency of first-person *and* third-person plural pronouns has been found in alt-right forums (Grover & Mark, 2019), as well as an increasing use of first-person and third-person plural pronouns related to group identification in online xenophobic forums (Bäck et al., 2018; Shrestha et al., 2020). When trying to detect xenophobic narratives in alt-right media as compared to regular conventional media, third-person plural pronouns were also significantly more frequent in the former than in the latter (Kaati et al., 2016b). In summary, while the use of first-person plural may have several different functions (see Pennebaker, 2011), it would seem that frequent use of third-person plural (*they*, *them*, etc.) among a group suggests that such a group is defining itself by the existence of an oppositional group—it would be an indicator of a strong feeling of group identity based on the existence of a contrasting out-group (Pennebaker, 2011; Pennebaker & Chung, 2012; Shrestha et al., 2020).

Moreover, research using LIWC to explore the use of language in political debate has revealed systematic differences both in semantics and in stylistic indices computed from function words such as articles, prepositions, or pronouns that are associated not only with ideological orientations (e.g., Sterling et al., 2020), but also with long-term trends of evolution of argumentation and reasoning as related to increasing polarization in the political game (e.g., Jordan et al., 2019), thus shedding some light on the relation that political reasoning styles keep with both ideology and historical context. While some previous research suggests that liberals favor analytical or systematic thinking and conservatives favor narrative or heuristic thinking (e.g., Jost & Krochik, 2014), there are also suggestions that political simple-mindedness might be more common in the opposition than in office (cf. Houck et al., 2020, for a review; and Conway and Zubrod, 2022, for an analysis of the relation between different aspects of cognitive complexity in the political realm). However, to our knowledge, there has been no specific exploration of how the condition of incumbent or challenger party might affect political language in reasoning and argumentation. This is partly because much of this research has regarded variables like reasoning style or perception of power (Körner et al., 2022) from a more individualistic perspective, as an enduring disposition of individuals or categories of people, rather than framing the issue within a more dynamic intergroup context, although there is also increasing evidence that the interpersonal or intergroup nature of the immediate context, as well as group relative power or status positions, are crucial determinants of intergroup language use and argumentation styles (Van Swol & Kane, 2019).

While it is true that there is no single LIWC category that readily captures intergroup polarization, nevertheless previous research on ideological radicalization suggests that emotional language (Pennebaker & Chung, 2008), reduced cognitive complexity and use of pronouns might be used as linguistic proxies for outgroup negativity (Shrestha et al., 2020). In particular, linguistic expression of negative emotions, markers of political simple-mindedness and a higher rate of self-exclusive pronouns should concur with higher levels of political polarization.

Moreover, most of this research on language use has analyzed the performance of English speakers (the language that LIWC's design is best fitted for), and therefore its extension to languages other than English may require significant adaptations, particularly when tackling the use of grammatical persons in a pronoun-drop language such as Spanish.

Research Overview and Hypotheses

The aim of this study is to test if political position (government vs. opposition) moderates the effect of ideology on both stereotypical framing of immigrants [RQ1] and the use of subtler language markers pointing to intergroup polarization when talking about immigration [RQ2]. In particular,

- (i) According to literature on immigration politicization (cf. Van der Brug et al., 2015), it would be expected that the *political position of challenger would lead to the accentuation of the opposition party's preferred framing of immigration* (i.e., *threat* for the right, *victim* for the left), thus reducing ambiguity on their stance and fostering polarization (i.e., increasing "politicization" of the issue) [H1].
- (ii) Drawing upon literature on intergroup conflict and language use (e.g., Prins, 2018; Shrestha et al., 2020), it would also be expected that increases in political antagonism and outgroup negativity, as reflected in the use of more self-exclusive pronouns (*you, they*), reduced cognitive complexity and more negative emotional language, would be associated to the political position of challenger. However, as right-wing parties seem more likely to weaponize the issue of immigration for political-electoral purposes (Odmalm & Bale, 2015; Odmalm, 2018), *an interaction effect of political position and ideology would be expected on those variables rather than a main effect of political position* [H2].
- (iii) Finally, given that the time period under analysis is long and the events in 2006 marked a turning point in the salience of the immigration issue both in public opinion and in parliamentary speeches in Spain, the adequacy of conducting a global analysis for the entire period or a period-segmented one was explored in our study, to check for the *possibility that politicization indicators mentioned in hypotheses 1 and 2 would increase after the "cayucos" crisis* [H3].

Method

Data

ParlSpeech, a database containing all the transcriptions of the Spanish Parliament from 1996 onwards, was accessed. The period compressed between 1996 and 2016 was selected, so as to cover ten years before key year 2006 and ten years later. Within

that time period, all the parliamentary speeches delivered in Spanish Congress of Deputies and containing at least one of a list of keywords related to immigration were selected (see appendix 1 for the keyword list). However, even if up to 15 different political parties were represented in Spanish Congress during the above period, only two of them had been alternating in Spanish government until 2016 and were therefore codable as “government party” and “main challenger party”: conservative Partido Popular was in power from 1996 to 2004, and again from 2011 to 2016, while moderate-left Partido Socialista Obrero Español was in government from 2004 to 2011. Therefore, only speeches delivered by a representative of one of those two parties were selected. As a result of this process, our dataset comprises 2,516 speeches delivered by conservative Popular Party (811 while in office and 373 in the opposition) and moderate-left Socialist Party representatives (736 while in office and 596 in the opposition) between 1996 and 2016. As all the speeches are coded by date, it can be identified whether the party of the politician delivering it was in government or in the opposition at the time of the speech delivery (see also some background demographic information of speakers in Appendix 2 in Supplemental Material).

Procedure

In order to assess to what extent each speech stereotypically frames immigrants as victims or as a threat, a semi-automatic procedure relying on a Gold Standard Corpus (GSC) of immigrant stereotyping (Sánchez-Junquera et al., 2021) was devised³. That GSC was constructed by manual annotation and resulted in a set of 1,479 sentences that were tagged as “threat-framing” sentences or as “victim-framing” sentences (see Appendix 3 in Supplemental Material and Sánchez-Junquera et al., 2021 for a more detailed description of the construction of this GSC, as well as for evidence of validity). In this paper, that GSC of threat-focused or victim-focused sentences was used to assign a “victim score” and a “threat score” to each of the 2,516 parliamentary speeches that constitute our database following the procedure explained below.

In order to select words that best help discriminate between both classes (i.e., the “victims” frame or the “threat” frame), a concept from information theory named Information Gain⁴ (IG, Larose, 2014) was applied. Based on their frequency distribution in the given classes in the GSC (“threat-framing” sentences or “victim-framing” sentences) IG assigned a score to each word according to its power to discriminate between such classes. Once the IG scores had been assigned to all the 7,269 different words contained in the 1,479 sentences of Sánchez-Junquera et al. (2021) GSC, they were sorted in descending order and the first 200 words were selected. Of these 200 words, 40 words were discarded for one of the following reasons: (1) they were the proper noun of a politician or a territory (e.g., Zapatero or Canary Islands), (2) it was a noun designating a month (e.g., April), (3) they had a very general meaning hardly associable to the semantic field of immigration (e.g., the verb to be), (4) they were words that are specific to parliamentary debate (e.g., deputy), or (5) they were function words (e.g., pronouns like “we” or “you”) that are the object of specific computation in this study.

Such a discarding process resulted in a dictionary of 160 words that best discriminated between “threat-framing” and “victim-framing” sentences. Then, a “victim” value and a “threat” value was assigned to each of these words depending on their relative frequency in the two classes that the 1,479 sentences of the GSC covered. For example, if a word appeared 30% of the time in sentences tagged with the frame “victim” and 70% of the time in sentences tagged with the frame “threat,” this particular word got a “victim” value of 30 and a “threat” value of 70. The last step of the procedure was to score each of our 2,516 speeches with both a “victim” score and a “threat” score that were the result of the addition of the “victim” and “threat” values of the words from the dictionary that appeared in each speech. The list of words that make up the dictionary with their corresponding “victim” and “threat” values and the list of discarded words can be found in additional documentation (see Appendix 4 in Supplemental Material).

Moreover, in order to analyze language style, the Spanish version (Ramírez-Esparza et al., 2007) of LIWC (Pennebaker et al., 2007) was used. When LIWC is used to elaborate an index, a score for each LIWC category is computed to reflect the percentage of that particular category over the total of words that were recognized by LIWC in this speech. LIWC is able to identify all personal pronoun categories, and it has been widely used to construct linguistic indexes aimed at detecting the use of certain grammatical persons in a text. However, Spanish -unlike English- is a pronoun drop language, and that means that the simple recount of personal pronouns that LIWC performs would miss most instances of use of grammatical persons of interest. A system that can also identify verbs that are conjugated in a certain grammatical person even when the corresponding pronoun is dropped would be required to accurately reflect the use of different grammatical persons in a speech that is delivered in Spanish. Therefore, another natural language processing tool, spaCy (Srinivasa-Desikan, 2018), able to identify all the grammatical categories in a speech, was used to construct those linguistic indexes aimed at detecting the use of certain grammatical persons. When the part-of-speech tagging of spaCy is used to elaborate an index, a score for each grammatical category is computed to reflect the percentage of that particular category over the total number of words in this speech.

Measures

Immigration stereotypical framing index. An index of immigration stereotypical framing was constructed with the score of the speech in the *victim frame* less the score of the speech in the *threat frame* divided by the number of words in the speech and multiplied by 100. In the Immigration stereotypical framing index, positive scores indicate that the immigrant group is presented more as a *victim* and negative scores indicate that the immigrant group is presented more as a *threat*.

Adversarial index. Using the morphology information provided by the tagger spaCy v3.0, the number of both singular and plural first-person and second-person pronouns and verbs in every speech was identified. The Adversarial index is computed as the number of second-person pronouns and verbs less the value of first-person pronouns and verbs divided by the number of words in the speech and multiplied by 100. In

this index, positive scores indicate that the speech focuses more on the political adversary and negative scores indicate that the speech focuses more on the speaker or their group. As such, it reflects how much the speech is anchored in the political adversary's actions and ideas rather than in the speaker's and their own party's actions and ideas. The adversarial genre has been studied in parliamentary and election debates, both in British (Bull & Wells, 2012) and Spanish contexts (Blas-Arroyo, 2003), but those pieces of research did not use natural language processing tools.

Ingroup vs outgroup index. Also using the morphology information provided by the tagger spaCy v3.0, the number of first-person plural pronouns and verbs in every speech was identified, and the resulting variable was named "ingroup." A similar process was carried out with third-person plural pronouns and verbs, and the resulting variable was named "outgroup." A novel index was created with the value of the "ingroup" variable less the value of the "outgroup" variable divided by the number of words in the speech and multiplied by 100. In this index, positive scores indicate that the speech refers more to the ingroup and negative scores indicate that the speech refers more to the outgroup. As noted in the introduction, previous work shows a consistent relation between use of pronouns and ideology (see Pennebaker & Chung, 2008; Seyle & Pennebaker, 2004; Grover & Mark, 2019; Bäck et al., 2018). But all those pieces of research computed the frequency of first-person and third-person plural words independently, and none of them used a relational index to measure the relative emphasis placed on the ingroup or the outgroup for discourses building on an intergroup context.

Analytic thinking index. Applying LIWC2007, the Analytic thinking index (Pennebaker et al., 2014) was computed to measure the use of a more analytic language or a more intuitive one. Initially labeled as the categorical-dynamic index, this index was later renamed analytic thinking in LIWC2015. The Analytic thinking index is computed as a simple algorithm: article + preposition – personal pronoun – impersonal pronoun – auxiliary verb – conjunction – adverb – negation (all categories computed as percentages over the total word count). Positive values in this index express more analytic thinking and negative values more intuitive thinking. This index has been extensively used in the long-term analysis of all kinds of political texts (Jordan et al., 2019), in electoral campaigns (Körner et al., 2022), and in social media messages (Hu & Kearney, 2021).

Positive vs negative emotion index. Applying LIWC 2007, a new emotion index was computed with the percentage of words expressing positive emotions less the percentage of words expressing negative emotions in each speech. In this index, positive scores indicate that the speech is mainly expressing positive emotions and negative scores indicate that the speech is mainly expressing negative emotions (see Appendix 5 in Supplemental Material for examples of speeches that scored high or low in all the previously described measures).

Party ideology. Speeches delivered by speakers belonging to conservative Popular Party were coded as PP, and speeches delivered by speakers belonging to moderate-left Partido Socialista Obrero Español were coded as PSOE.

Political position of the speaker. Based on the date of the speech and the speaker's party, each speech was coded as "government" if it was delivered by a speaker whose

party was in power at the time of the speech delivery or ‘opposition’ if it was delivered by a speaker whose party was the main challenger at the time of the speech delivery.

Historical period. In order to test if partisan strategies about immigration changed in relation with the “canoe crisis” in 2006, speeches delivered between 1996 and 2005 were coded as “period 1” and speeches delivered from 2006 onwards were coded as “period 2.”

An observational design was used with three between subjects variables: ideology (PP vs PSOE), political position (government vs opposition) and historical period: 1996–2005, and 2006–2016.

Statistical Analyses

Our main objective was to evaluate the effect of the interaction between ideology and political position in the two historical periods considered (1996–2005, and 2006–2016). As a first step, with the aim of testing if the above interaction could be moderated by the historical period variable, a three-factor (ideology x political position x historical period) analysis of variance (ANOVA) was computed on the immigration stereotypical framing index and all the linguistic indices. When the triple interaction was statistically significant, the effects of the two independent variables (ideology and political position) and their interaction were examined with separate ANOVAs for the data from each historical period. When the triple interaction was not statistically significant, the effects of the two independent variables (ideology and political position) and their interaction were examined in the whole dataset, merging the data from the two historical periods. A statistically significant interaction effect between both factors was followed by a test of simple effects, and 95% CI for mean differences. For all the analysis, we also report partial eta-squared (η_p^2) considering cut-offs of 0.01, 0.06, and 0.15 for small, medium, and large effects, respectively (Cohen, 1988).

Kolmogorov-Smirnov test indicated violation of the assumption on normality for all of the measures, however we decided to proceed with parametric analysis, given that research shows the robustness of ANOVA to violations of normality (Blanca et al., 2017; Schmider et al., 2010). Statistical significance was set by convention at $p < 0.05$. For pairwise comparisons, 95% confidence intervals are reported. All statistical analyses were conducted using SPSS Statistics (version 28.0; IBM Corp., Armonk, NY, USA).

Results

The triple interaction (ideology x political position x historical period) was statistically significant for the variables: Ingroup vs outgroup index, and Analytic thinking index. So, for these dependent variables, the results of the two-factor (ideology x political position) between-subjects ANOVA are presented separately for each historical period in Table 1.

The above triple interaction was not statistically significant for the following variables: Immigration stereotypical framing index, Adversarial index, and Positive vs. negative emotion index. Consequently, for these dependent variables, the two-factor

(ideology x political position) ANOVA for the whole dataset (1996–2016) are presented in Table 2.

Interaction of Ideology and Political Position on Stereotyping of Immigrants

For the Immigration stereotypical framing index, the triple interaction of ideology x political position x historical period is not significant, but a significant two factor interaction between ideology and political position is observed for the whole period ($p < .001$) (see Table 2). The simple effects analysis (see Figure 2) shows that both, PSOE and PP, change their positions as result of being in government or in the opposition ($p < .001$): PSOE uses more victim markers in the opposition than in government (mean difference = -37.09 , 95% CI: -48.43 , -25.75); and, on the other hand, PP shows more victim markers in government than in the opposition (mean difference = 56.16 , 95% CI: 43.30 , 69.04). Moreover, both parties differ when they are in government and when they are in the opposition ($p < .001$): in government, PP uses more victim markers than PSOE (mean difference = 21.48 , 95% CI: 11 , 31.95); while, in the opposition, PSOE uses more victim markers than PP (mean difference = -71.78 , 95% CI: -85.36 , -58.19).

Interaction Effects on Language Style

For the Adversarial index, the three-factor interaction (ideology x political position x historical period) is not statistically significant, but a significant two factor interaction

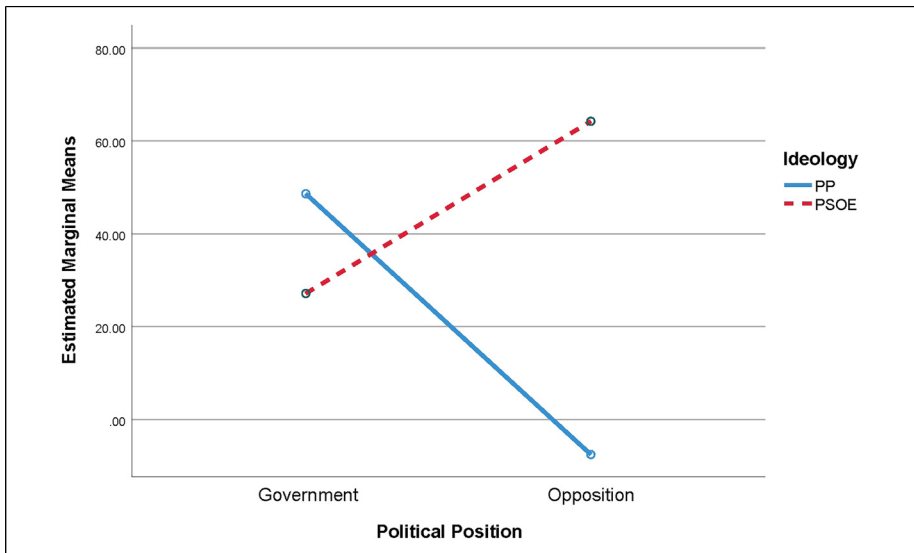


Figure 2. Immigration stereotypical framing index by ideology and political position (1996–2016).

Table 1. Two-Factor ANOVA on the Variables: Ingroup vs Outgroup Index and Analytic Thinking Index.

Index	Effects	First period (1996–2005)					Second period (2006–2016)				
		df	F	P	η_p^2		df	F	p	η_p^2	
Ingroup vs outgroup	Ideology	1, 1268	1.41	.236	.001		1, 1240	0.21	.645	.000	
	Political position		0.26	.610	.000			13.75	<.001***	.011	
Analytic thinking	Ideology x Political position		0.16	.694	.000			18.70	<.001***	.015	
	Ideology	1, 1268	28.73	<.001***	.022		1, 1240	3.81	.051	.003	
	Political position		68.26	<.001***	.051			20.62	<.001***	.016	
	Ideology x Political position		10.62	.001***	.008			5.67	.017*	.005	

Note: * $p < .05$. ** $p < .01$. *** $p < .001$.

between ideology and political position is observed in the whole dataset ($p < .001$) (see Table 2). The simple effects analysis for the Adversarial index (see Figure 3) shows that conservative PP is more adversarial in the opposition than in government ($p < .001$), while moderate-left PSOE doesn't change their style as a result of being in government or in the opposition ($p = .237$). Moreover, both PP and PSOE differ in this linguistic index when they are in government ($p < .001$) and when they are in the opposition ($p = .002$): when in government, PP uses a less adversarial style than PSOE (mean difference = $-.77$, 95% CI: $-.97, -.57$) while, in the opposition, PP uses a more adversarial style than PSOE (mean difference = $.34$, 95% CI: $.13, .56$).

As for the Ingroup vs outgroup index, the three-factor interaction (ideology $\times$ political position $\times$ historical period) is statistically significant, $F(1, 2508) = 10.88$, $p < .001$, $\eta_p^2 = .004$. Therefore, the two-factor interaction ideology $\times$ political position was examined in each historical period (see Table 1) and it was observed that this interaction is statistically significant in the second historical period ($p < .001$) but not in the first one ($p = .694$). In the first period, neither the interaction effect nor the main effects of party or political position are statistically significant (see Figure 4a). As for the second period (see Figure 4b), moderate-left PSOE does not change their ingroup-outgroup emphasis when their political position changes ($p = .656$) but conservative PP does ($p < .001$), using more outgroup markers in the opposition than in government (mean difference = -1202.17 , 95% CI: $-1626.08, -778.26$). Moreover, PP uses more ingroup markers in government than PSOE ($p < .001$) (mean difference = 716.24 , 95% CI: $339.86, 1092.61$) and more outgroup markers in the opposition than PSOE ($p = .01$) (mean difference = -578.23 , 95% CI: $-1,029.05, -127.40$).

Regarding the Analytic thinking index, the three-factor interaction (ideology $\times$ political position $\times$ time period) is statistically significant, $F(1, 2508) = 11.08$, $p < .001$, $\eta_p^2 = .006$. The analysis of the two-factor interaction (ideology $\times$ political position) in each historical period shows that this interaction is statistically significant in both periods ($p < .001$ and $p = .017$, respectively) (see Table 1). In the first historical period (see Figure 5a), the simple effects analysis shows that both parties are less analytic in the opposition than in government ($p < .001$) (mean difference = $-.68$, 95% CI: $-.87, -.50$ for PP and mean difference = $-.30$, 95% CI: $-.44, -.16$ for PSOE). Moreover, PP and PSOE do not significantly differ when in government ($p = .063$), but they do when they are in the opposition ($p < .001$), so that PP is less analytic than PSOE (mean difference = $-.51$, 95% CI: $-.70, -.32$).

In the second historical period (see Figure 5b), the simple effects analysis shows that only moderate-left PSOE changes ($p < .001$) when their political position changes (mean difference = $-.34$, 95% CI: $-.47, -.20$), being less analytic in the opposition than in government. Conservative PP does not change ($p = .135$) when they are in government or in the opposition. Moreover, in this historical period, PP and PSOE significantly differ when in government ($p < .001$) but not in the opposition ($p = .780$), so that, when in government, PSOE is more analytic than PP (mean difference = $.21$, 95% CI: $.09, .33$).

Finally, on the Positive vs negative emotion index, the three-factor interaction (ideology $\times$ political position $\times$ historical period) is not statistically significant.

Table 2. Two Factor ANOVA on the Variables: Immigration Stereotypical Index, Adversarial Index and Positive vs. Negative Emotion index.

Index	Effects	1996–2016			
		df	F	p	η^2_p
Immigration stereotypical framing	Ideology	1, 2512	33.06	<.001***	.013
	Political position		4.75	.029*	.002
	Ideology x Political position		113.62	<.001***	.043
	Ideology	1, 2512	0.03	.867	.000
Adversarial	Political position		40.34	<.001***	.016
	Ideology x Political position		22.91	<.001***	.009
Positive vs negative emotion	Ideology	1, 2512	7.80	.005**	.003
	Political position		197.72	<.001***	.073
	Ideology x Political position		3.71	.054	.001

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

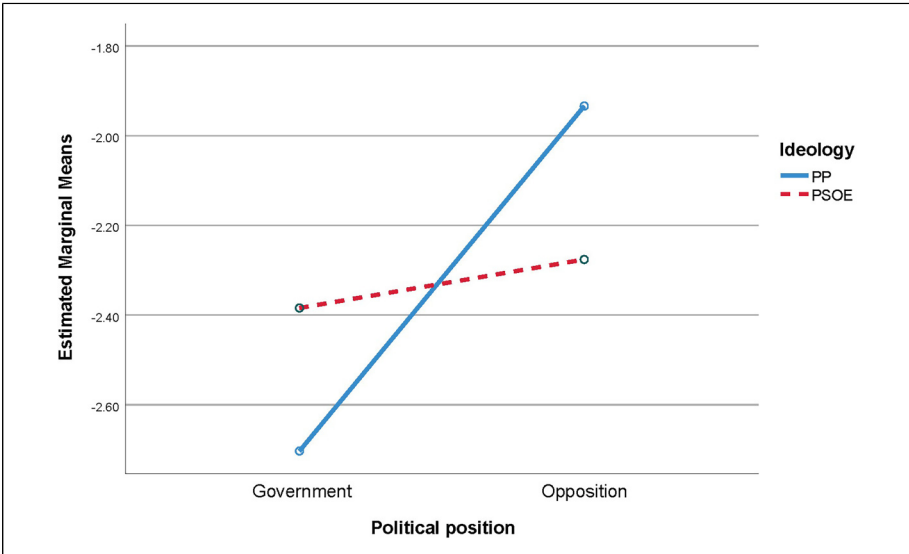


Figure 3. Adversarial index by ideology and political position (1996–2016).

The two-factor interaction is only marginally significant in the whole dataset ($p = .054$) and the two main effects of ideology and political position are statistically significant ($p = .005$ and $p < .001$, respectively) (see Table 1). The strongest effect is for political position: both parties use a more positive emotional language being in government ($M = 0.021$) than in the opposition ($M = 0.011$) (mean difference = $-.009$, 95% CI: $-.011, -.008$). Regarding the effect of ideology, moderate-left PSOE conveys more positive emotions ($M = .017$) than conservative PP ($M = .015$) (mean difference = $-.002$, 95% CI: $-.003, -.001$) in their speeches about immigration.

Discussion

Effects of Ideology and Political Position on Language Use Regarding Immigrants

As it has been repeatedly shown in previous research, content of immigrants' representations is dependent on political ideology (Benson, 2013; Simmons et al., 2018; Mendelsohn et al., 2021; Kocijan & Kukec, 2022). However, our results show an interaction between ideology and political position on particular frames to stereotype immigrants over the whole 20-year period considered [RQ1]. This result suggests that immigration is not only an ideologically loaded topic, but also an issue that has been used in partisan battles in Spanish parliamentary debate in the last decades. Politicization, as discussed in the introduction to this paper, involves further polarization of positions due to strategic considerations in party competition, and this is exactly what our results suggest.

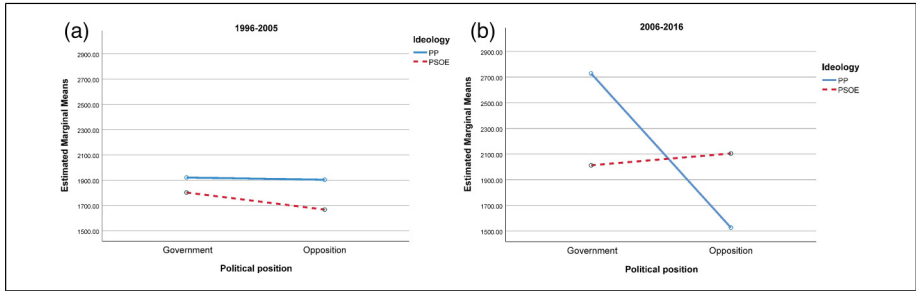


Figure 4. 4a and 4b. Ingroup and outgroup index by ideology and political position (1996–2005) and (2006–2016).

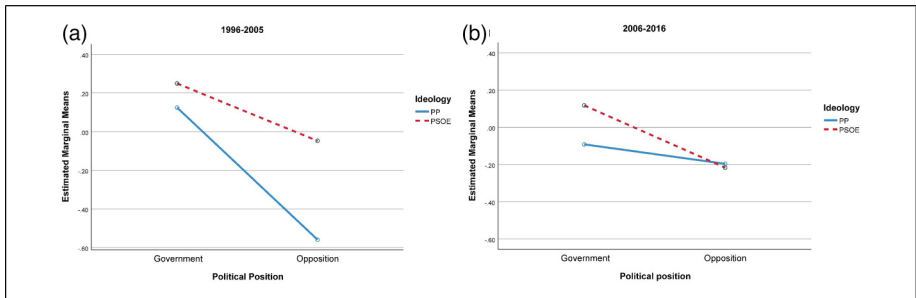


Figure 5. 5a and 5b. Analytic index by ideology and political position (1996–2005) and (2006–2016).

The ANOVA performed on the Immigration stereotypical framing index yields a significant interaction effect between party ideology and political position, showing that both main parties in Spanish parliament accentuate their preferred stereotypical representation of immigrants (a *threat frame* for conservative PP and a *victim frame* for moderate-left PSOE) when they are in the opposition, thus polarizing positions on the issue [H1]. That such polarization happens in a topic such as immigration is not surprising, since issues that align with pre-existing ideological differences are particularly susceptible to politicization (Berkhout & Van der Brug, 2013).

On the other hand, both parties moderate their typical positions when they are in office to the point that conservative PP in government emphasizes a victim frame more than moderate-left PSOE. This unexpected result evidences the weight of political position when accounting for parties' stances on immigration and should be tested across other controversial issues.

In order to explore in more depth the intergroup nature of the context in which politicians speak about immigrants, an Adversarial index was created that examines the relative emphasis on second person pronouns as compared to self-inclusive pronouns. It should be noted that parliamentary exchanges are a genre usually characterized for an

intensive use of second person markers in a discursive strategy that blames the political adversary for their words and actions, and it has already been noted how surprising it is that this feature has not been more used to explore parliamentary conflict.

Our significant two-factor interaction between ideology and political position on the Adversarial index in the whole period indicates that, in their speeches about immigrants, conservative PP makes a very different use of adversarial markers when their political position changes but moderate-left PSOE does not change significantly. This would suggest that, in the subtler plane of interactional dynamics that pronoun-based indices capture, conservative PP is more sensitive to strategic considerations or potential electoral gain when speaking about immigration [H2]. This would lead them to accentuate antagonism when they act as challengers to the party in office, but to try and restrain confrontation over controversial issues when they regain power.

On the other hand, our second pronoun-based index, the ingroup vs outgroup index, tackles a subtler layer of representation of the relations between natives and immigrants than the immigration stereotypical framing index, more focused on content, can capture. As noted in the introduction to this paper, the presence of third-person plural pronouns has been found to predict xenophobic discourse (Seyle & Pennebaker, 2004; Kaati et al., 2016b) and high rates of “they” pronouns have been associated in past research with an increased intergroup bias to separate one collective from another (Gustafsson Sendén et al., 2014; Perdue et al., 1990). However, research also shows that majority groups (Van Swol & Carlson, 2015), high-status groups (Pennebaker & Chung, 2012), authoritarian leaders (Liu, 2022), and people with stronger social identities (Bäck et al., 2018; Shrestha et al., 2020) tend to use more *both* first person *and* third person *plural* pronouns. Some experts have suggested that, in fact, a greater use of all kinds of *plural* (as compared to singular) pronouns would be indicative of a general “outward focus” (Pennebaker & Chung, 2012), postulated as a maker of high status and strong leadership (Kacwicz et al., 2014). This mixed evidence makes it difficult to decide if a higher rate of third-person plural pronouns specifically reflects negative othering or such more general ‘outward focus’. However, our Ingroup vs. outgroup index is a relational one that reflects a greater *relative* emphasis on the outgroup rather than on the ingroup (or vice versa). Therefore, our results cannot be regarded as indicative of a greater general “outward focus” but they are more likely to be reflecting a more polarized debate. Polarization in a debate tends to organize individuals and groups around exclusive identities, and it has been found to be associated with the use of less self-inclusive and more self-exclusive pronouns (Prins, 2018). Our relational index reflects an emphasis on self-exclusive pronouns and verbs, and therefore may be interpreted as an indicator of polarization in political debate, making use of an *us vs them* rhetorical strategy.

The triple interaction found in the Ingroup vs outgroup index indicates that the joint effect of ideology and political position does not occur in the first historical period (1996–2005) but only in the second period, as H3 would suggest, and it applies only to conservative PP, thus confirming H2. In the second historical period, the right-wing party is anchored in ingroup markers being in the government but in

outgroup markers being in the opposition, while the left-wing party does not change. Therefore, this would suggest that it is essentially the political right that has changed their discursive strategy related to immigration in the last decades, using this topic in a polarization strategy. Again more research is needed to see if the use of a more polarized narrative in recent times is a discursive strategy used by the political right only on immigration or it also occurs in general with all controversial political issues. In the same way, future research should also confirm if the political left reproduces identical linguistic patterns in issues other than immigration.

Regarding the Analytic thinking index, the three-factor interaction shows that the two mainstream parties in the Spanish Parliament modulate the complexity of their discourse on immigration according to their political position and the historical context. In the first period, as expected, both parties are less analytic in the opposition than in government, but PP is less analytic than PSOE in the opposition. However, in the second period only PSOE changes when their political position shifts, adopting a more simplifying discourse. Therefore, our results would not entirely confirm H2 or H3. Even though they point to the relevance of the independent variables involved in both hypotheses, rather our findings suggest that analytic thinking is influenced by a complex interaction between political ideology, party competition, and historical context that deserves closer inspection.

As for the Positive vs negative emotional index, it is not surprising that both parties show a more positive tone in their discourse when being in government. It is not surprising either that ideology also has a (smaller) main effect on this emotional index, considering that right-wing parties tend to present immigrants as more of a threat than left-wing ones. However, no interaction effect was found, as H2 would have predicted.

Immigration in the Context of Partisan Battle

Globally considered, then, our results suggest that immigration is an issue that had been politicized in Spanish parliamentary debate by mainstream parties in the last decades, even before any radical-right party appeared in the Spanish political scenario.

Results on the preferred framing for immigrants of both parties reveal polarizing stances on the issue when both of them find themselves in the challenger position, as might be expected [RQ1]. However, it is not only the content of the characterization of immigrants that changes when there is a shift in political positions, but also some subtler elements of language usage [RQ2], and particularly the nature of intergroup language as revealed by the use of grammatical persons. Even if our results show that both parties tend to appeal to more negative emotions and less analytic thinking when they are in the opposition than when they are in government, the analyses performed on the indices based in the use of grammatical persons show that it is only conservative PP that uses more self-exclusive pronouns and verbs when they are in the opposition, thus building up intergroup differentiation and conflict.

Both our Ingroup vs. outgroup index and our Adversarial index can be interpreted as markers of intergroup differentiation and conflict, although of a different nature in both

cases. The Adversarial index, exploring the relative emphasis on self-exclusive “you” words as compared to self-inclusive pronouns, reflects conflict and polarization between political parties. And indeed the fact that self-exclusive “you” words are used at a higher rate to discuss immigration by conservative Popular Party when they are in the opposition would suggest that they are using this topic to accentuate confrontation in partisan battle.

The Ingroup vs. outgroup index, on the other hand, reflects a more or less divisive representation of the relation between natives and immigrants, as it has been argued. And research has shown that people unconsciously tend to rate more positively whatever is paired with a first-person plural than whatever is paired with a third-person plural, so that a higher proportion of outgroup designators (“they,” “them”) will automatically provide the speech with a more negative evaluative tone, thus perpetuating intergroup biases (Perdue et al., 1990). In fact, people actively generate such evaluative associations by choosing more positive semantic contexts for self- versus other pronouns, thus producing and reproducing differences between social categories in evaluative associations (Gustafsson Sendén et al., 2014). The evaluative gap between self-inclusive and self-exclusive pronouns was even shown to be magnified (i) in the case of collective rather than individual pronouns, and (ii) in an intergroup condition versus an individual or interpersonal condition (Gustafsson Sendén et al., 2014). Our data meet both these magnifying conditions, because (i) the Ingroup vs. outgroup index deals with plural pronouns, and (ii) a parliamentary debate is a prototypical intergroup context in which partisan differences activate both self- and group-serving biases. Therefore, a higher rate of “they” words in conservative PP’s speeches when they are in the opposition suggests the party tends to accentuate negative “othering” of immigrant persons in order to regain power.

Moreover, the fact that the effects of party ideology and political position are partially concurrent and parallel on both intergroup conflict indices would suggest the possibility that one thing is used *in order to* achieve the other -i.e., a more divisive representation of natives and immigrants helps polarize partisan battle. Although that possibility cannot be tested with our present data on intergroup indices, our results nonetheless show that language patterns (even those of non-semantic function words such as grammatical person markers) can offer a lens into psychosocial processes that can provide rich data about political intergroup dynamics.

Finally, the fact that such intergroup-polarizing linguistic changes only happen in conservative PP, while moderate-left PSOE does not change intergroup language when shifting from government to the opposition, is quite consistent with political science literature arguing that mainstream right-wing parties, rather than only alt-right parties, have had a very significant role in the politicization of immigration issues in Europe (cf. Bale, 2008; Green-Pedersen & Krogstrup, 2008; Meyer & Rosenberg, 2015). Maybe Spain was never exceptional in the issue of politicization of immigration, after all (Amat & Garcés-Mascreñas, 2018; Morales et al., 2015; Ros & Morales, 2015), but politicization of the topic had been happening even in absence of radical right parties, and that helps explain the very successful irruption of Vox shortly after the end of the period covered by our data (cf. Turnbull-Dugarte, 2019).

Limitations and Future Research

This work presents two key contributions: one regarding the methodology employed and the other offering a promising insight into the understanding of immigration in the context of partisan battle. As for the methodological one, four linguistic indices have been created ex-novo—the Immigrant stereotypical framing index, the Ingroup vs outgroup index, the Adversarial style index, and the Positive vs negative emotion index. The construction of the first index involved human annotation of a corpus. To the best of our knowledge, it is also the first time that the Information Gain methodology has been used to expand human annotation from a small corpus to a larger one. The other three new indexes can be applied to any Spanish text using available resources such as the spaCy v3.0 tagger and the LIWC2007 dictionary. Additionally, these contributions are also relevant because they explore the adaptations required by NLP tools to Spanish, an underrepresented language in the emerging literature of computational social science.

However, the behavior of these novel indexes should be further investigated and improved in future research. First and foremost, a closer examination of the specific referents of pronoun-based indexes in every instance of use is called for. Although parliamentary practice makes it sensible to assume that *you* words may be the main resource to signal intergroup distance with political adversaries, while *they* words may operate as markers of intergroup distance with the social categories the speech is about (immigrants, in this case), the truth is that this is just a fair assumption due to the exploratory nature of this piece of research. Therefore, our pronoun-based indexes should be refined to be able to accurately identify their referents—a shortcoming that is shared with most research using dictionary-based tools such as LIWC. Moreover, the behavior of such indexes in this and other contexts should be further examined. Just to give some examples, the way intergroup conflict indexes behave in communicative interactions in which the interlocutor is not present—such as political speeches at rallies, press conferences or opinion articles—might be explored; and the positivity or negativity of a discourse should be evaluated not only by relying on words that express emotions, but also on other linguistics markers.

Regarding the study of immigration in the context of partisan battle, the main contribution of this research is to show that political position, linked to intergroup conflict, is a key variable. The political ideology of the speaker is not enough to explain the variability that appears in the content and linguistic style of parliamentary discourses about immigration.

The fact that these linguistic indices are sensitive to the dynamics of political battle opens up new avenues for studying how stereotypes are constructed in everyday language by political elites or other population groups who are difficult to study with other methodologies. Future research should examine whether the rhetorical strategies identified by these linguistic indexes are also present in other controversial issues, for example, women's rights, LGBTQ+ communities, climate change, etc.

On the other hand, the objectives of this research made it appropriate to focus only on parties that have held government responsibilities, specifically majority parties, as they are the ones who truly feel themselves to be the main challenger when they are not in power. However, a thorough analysis of politicization of this and other issues in Spain would require to examine whether the same results apply to the discourse of minority

parties in the Spanish parliament. Some of these political actors combine their ideological affiliation on the right-left axis with the defense of territorial interests linked to different regions of Spain, making the group dynamics more complex and presenting a greater challenge in the operationalization of indices such as the Ingroup vs outgroup index.

Moreover, from 2016 until now, the landscape of the Spanish Parliament has experienced a deep transformation with the entry of a radical right party in Congress. Immigration is at the center of radical-right political discourse and it would be necessary to examine if these indices are useful to detect ideological differences and the dynamics of intergroup conflict in this new context. Also, it could be interesting to measure if and how the arrival of the radical-right parties has affected the communication patterns of the rest of political actors in the Spanish Parliament.

Conclusions

Natural language processing (NLP) opens up many promising possibilities of psychosocial research in fields and topics in which targets of interest were difficult to reach through traditional methods. This has been typically the case with political elites, although their role in shaping public opinion and discourse on issues that are essential for civic life has long been commented upon (Helbling, 2014). This paper has tried to exploit such possibilities to tentatively explore the role of Spanish political elites in the politicization of immigration, one of the issues (politicization, rather than immigration) that has recently posed a biggest threat to democratic systems in Europe. Inasmuch as this kind of research can illuminate dynamics that deeply affect peaceful coexistence, it would be important that social and computer scientists joined forces to advance and refine reliable psychosocial indicators that leverage the possibilities for social understanding that NLP offers, and that they did so in languages other than English. This piece of research is a first step in that direction.

Acknowledgements

Alejandro Toselli is supported by a María Zambrano grant from the Spanish Ministerio de Universidades and the European Union NextGenerationEU/PRT. Berta Chulvi and Paolo Rosso are supported by FairTransNLP-Stereotypes PID2021-124361OB-C31 funded by MCIN/AEI/10.13039/501100011033 and by ERDF, EU A way of making Europe. We would like to thank Dr. Javier Sanchez-Junquera for assistance with information retrieval to build the corpus and the anonymous reviewers for their valuable comments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Berta Chulvi  <https://orcid.org/0000-0003-1169-0978>

Supplemental Material

Supplemental material for this article is available online.

Notes

- 1 Although Spain is a parliamentary system with up to 15 political parties represented in Congress, it essentially worked until 2018 as a two-party system, with two large nationwide parties alternating in government and forming single-party cabinets with the occasional parliamentary support of some small regional parties. Moderate-left Socialist Party was in office from 1982 to 1996, and then again from 2004 to 2011. Conservative Popular Party held office from 1996 to 2004, and then again from 2011 to 2018.
- 2 The only Spanish radical-right party represented in the Spanish parliament to this day, Vox, first emerged as a prime contender in regional elections in 2018 and national elections in 2019.
- 3 A Gold Standard Corpus (GSC) is a manually annotated corpus of text that, after due validation, can be used to train and evaluate the performance of automatic and semi-automatic classifiers in discrimination tasks (cf. Wissler et al., 2014). In this case, the GSC described and validated in Sánchez-Junquera et al. (2021) was used as a baseline upon which a semi-automatic procedure was devised to evaluate “victim-framing” and “threat-framing” in a new dataset comprising all immigration speeches delivered by PP and PSOE representatives in the Spanish Congress between 1996 and 2016.
- 4 Information gain (IG) is a method that helps identify which words or phrases in a text are the most informative or relevant to a particular topic or classification. It is often used in text classification tasks, such as determining if a piece of text is positive or negative, or identifying the topic of a text. The method calculates the reduction in entropy after a specific feature (word or phrase) is observed. Entropy is a measure of the uncertainty or randomness in a set of data. By reducing the uncertainty, the feature is considered more informative.

References

- Abdou, L. H., Bale, T., & Geddes, A. P. (2022). Centre-right parties and immigration in an era of politicization. *Journal of Ethnic and Migration Studies*, 48(2), 327–340. <https://doi.org/10.1080/1369183X.2020.1853901>
- Alonso, S., & Fonseca, S. C. (2011). Immigration, left and right. *Party Politics*, 18(6), 865–884. <https://doi.org/10.1177/1354068810393265>
- Alonso, S., & Rovira, C. (2015). Spain: No country for the populist radical right? *South European Society and Politics*, 20(1), 21–45. <https://doi.org/10.1080/13608746.2014.985448>
- Amat i Puigsech, D., & Garcés-Mascareñas, B. (2018). *Politicization of immigration in Spain: an exceptional case?*. http://ceaseval.eu/publications/14_AmatGarcés_WP5_Spain.pdf.
- Anolli, L., Zurloni, V., & Riva, G. (2006). Linguistic intergroup bias in political communication. *The Journal of General Psychology*, 133(3), 237–255. <https://doi.org/10.3200/GENP.133.3.237-255>

- Bäck, E. A., Bäck, H., Gustafsson Sendén, M., & Sikström, S. (2018). From I to we: Group formation and linguistic adaption in an online xenophobic forum. *Journal of Social and Political Psychology*, 6(1), 76–91. <https://doi.org/10.5964/jsp.p.v6i1.741>
- Bale, T. (2008). Turning around the telescope: Center-right parties and immigration and integration policy in Europe. *Journal of European Public Policy*, 15(3), 315–331. <https://doi.org/10.1080/13501760701847341>
- Benson, R. (2013). *Shaping immigration news: A French-American comparison*. Cambridge University Press.
- Berkhout, J., & Van der Brug, W. (2013). *Why do events only sometimes lead to politicization? Assessing the politicization of migration in Western Europe, 1995-2009*. ECPR General Conference, Bordeaux, France. <https://ecpr.eu/Events/Event/PaperDetails/2424>.
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 552–557. <https://doi.org/10.7334/psicothema2016.383>
- Blas-Arroyo, J. L. (2003). “Perdóneme que se lo diga, pero vuelve usted a faltar a la verdad, señor Gonzalez”: Form and function of politic verbal behaviour in face-to-face Spanish political debates. *Discourse & Society*, 14(4), 395–423. <https://doi.org/10.1177/095792650301400400>
- Brooks, C., Manza, J., & Cohen, E. D. (2016). Political ideology and immigrant acceptance. *Socius*, 2, 1–12. <https://doi.org/10.1177/2378023116668881>
- Bull, P., & Wells, P. (2012). Adversarial discourse in prime minister’s questions. *Journal of Language and Social Psychology*, 31(1), 30–48. [https://doi.org/10.1177/0261927\(11425034](https://doi.org/10.1177/0261927(11425034)
- Charlesworth, T. E., & Banaji, M. R. (2022). Word embeddings reveal social group attitudes and stereotypes in large language corpora. In M. Dehghani & R. L. Boyd (Eds.), *Handbook of language analysis in psychology* (pp. 494–408). Guilford Press.
- Cichocka, A., Bilewicz, M., Jost, J. T., Marrouch, N., & Witkowska, M. (2016). On the grammar of politics—or why conservatives prefer nouns. *Political Psychology*, 37(6), 799–815. <https://doi.org/10.1111/pops.12327>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum. <https://doi.org/10.4324/9780203771587>
- Consterdine, E., & Kintz, M. (2018). *State-of-the-art report on public attitudes, political discourses and media coverage on the arrival of refugees*. CEASEVAL. http://ceaseval.eu/publications/02_Consterdine_State-of-the-art_report_on_public_attitudes.pdf.
- Conway, L. G., & Zubrod, A. (2022). Are U.S. presidents becoming less rhetorically complex? Evaluating the integrative complexity of Joe Biden and Donald Trump in historical context. *Journal of Language and Social Psychology*, 41(5), 613–625. <https://doi.org/10.1177/0261927X221081126>
- Dennison, J., & Mendes, M. S. (2019). *When do populist radical right parties succeed? Salience, stigma, and the case of the end of Iberian ‘exceptionalism’*. European University Institute. <http://hdl.handle.net/1814/62226>.
- De Wilde, P. (2011). No polity for old politics? A framework for analyzing the politicization of European integration. *Journal of European Integration*, 33(5), 559–575. <https://doi.org/10.1080/07036337.2010.546849>
- Dovidio, J., Hewstone, M., Glick, P., & Esses, V. (2010). Prejudice, stereotyping and discrimination: Theoretical and empirical overview. In J. F. Dovidio, M. Hewstone, & P. Glick (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 3–28). Sage.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social perception: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>

- Formanowicz, M. (2020). Verb intergroup bias: Verbs are used more often in reference to in-groups than out-groups. *Social Psychological and Personality Science*, 11(6), 854–864. <https://doi.org/10.1177/1948550619893957>
- Gattinara, P. C., & Morales, L. (2017). The politicization and securitization of migration in Western Europe: Public opinion, political parties and the immigration issue. In P. Bourbeau (Ed.), *Handbook of migration and security* (pp. 273–294). Edward Elgar.
- Gaucher, D., Friesen, J. P., Neufeld, K. H. S., & Esses, V. M. (2018). Changes in the positivity of migrant stereotype content: How system-sanctioned pro-migrant ideology can affect public opinions of migrants. *Social Psychological and Personality Science*, 9(2), 223–233. <https://doi.org/10.1177/1948550617746463>
- Goffman, E. (1974). *Frame analysis*. Harper & Row.
- Gonzalez-Enriquez, C. (2017). *The Spanish exception: Unemployment, inequality and immigration, but no right-wing populist parties*. Real Instituto Elcano. <https://www.realinstitutoelcano.org/en/work-document/the-spanish-exception-unemployment-inequality-and-immigration-but-no-right-wing-populist-parties/>.
- Graf, S., Bilewicz, M., Finell, E., & Geschke, D. (2013). Nouns cut slices: Effects of linguistic forms on intergroup bias. *Journal of Language and Social Psychology*, 32(1), 62–83. <https://doi.org/10.1177/0261927X12463209>
- Grande, E., Schwarzbözl, T., & Fatke, M. (2019). Politicizing immigration in Western Europe. *Journal of European Public Policy*, 26(10), 1444–1463. <https://doi.org/10.1080/13501763.2018.1531909>
- Green-Pedersen, C., & Krogstrup, J. (2008). Immigration as a political issue in Denmark and Sweden. *European Journal of Political Research*, 47(5), 610–634. <https://doi.org/10.1111/j.1475-6765.2008.00777.x>
- Greussing, E., & Boomgaarden, H. G. (2017). Shifting the refugee narrative? An automated frame analysis of Europe's 2015 refugee crisis. *Journal of Ethnic and Migration Studies*, 43(11), 1749–1774. <https://doi.org/10.1080/1369183X.2017.1282813>
- Grover, T., & Mark, G. (2019). Detecting potential warning behaviours of ideological radicalization in an alt-right subreddit. *Proceedings of the Thirteenth International AAAI Conference on Web and Social Media (ICWSM 2019)*. <https://ojs.aaai.org/index.php/ICWSM/article/view/3221>.
- Gustafsson Sendén, M., Lindholm, T., & Sikström, S. (2014). Selection bias in choice of words: Evaluations of “I” and “We” differ between contexts, but “They” are always aorse. *Journal of Language and Social Psychology*, 33(1), 49–67. <https://doi.org/10.1177/0261927X13495856>
- Haas, H. G., & Natter, K. (2015). *The determinants of migration policies. Does the political orientation of governments matter?* International Migration Institute. <https://www.migrationinstitute.org/publications/the-determinants-of-migration-policies-does-the-political-orientation-of-governments-matter>.
- Heizmann, B., & Huth, N. (2021). Economic conditions and perceptions of immigrants as an economic threat in Europe: Temporal dynamics and mediating processes. *International Journal of Comparative Sociology*, 62(1), 56–82. <https://doi.org/10.1177/0020715221993529>
- Helbling, M. (2014). Framing immigration in Western Europe. *Journal of Ethnic and Migration Studies*, 40(1), 21–41. <https://doi.org/10.1080/1369183X.2013.830888>
- Hooghe, L., & Marks, G. (2009). A postfunctionalist theory of European integration: From permissive consensus to constraining dissensus. *British Journal of Political Science*, 39(1), 1–23. <https://doi.org/10.1017/S0007123408000409>

- Houck, S. C., Conway, L. G. III, & Zubrod, A. (2022). Automated integrative complexity: Language analysis tools for psychological research. In M. Dehghani & R. L. Boyd (Eds.), *Handbook of language analysis in psychology* (pp. 173–188). Guilford Press.
- Houck, S. C., Huber, K. J., Ess, M., & Proulx, M. L. (2020). Cognitive complexity in political contexts. *Oxford research encyclopedia of politics*. Published online: <https://doi.org/10.1093/acrefore/9780190228637.013.1024>
- Hu, L., & Kearney, M. W. (2021). Gendered tweets: Computational text analysis of gender differences in political discussion on Twitter. *Journal of Language and Social Psychology*, 40(4), 482–503. <https://doi.org/10.1177/0261927X20969752>
- Jordan, K. N. (2022). Language analysis in political psychology. In M. Dehghani & R. L. Boyd (Eds.), *Handbook of language analysis in psychology* (pp. 159–172). Guilford Press.
- Jordan, K. N., Sterling, J., Pennebaker, J. W., & Boyd, R. L. (2019). Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proceedings of the National Academy of Sciences*, 116(9), 3476–3481. <https://doi.org/10.1073/pnas.1811987116>
- Jost, J. T., & Krochik, M. (2014). Ideological differences in epistemic motivation: Implications for attitude structure, depth of information processing, susceptibility to persuasion, and stereotyping. In A. J. Elliot (Ed.), *Advances in motivation science* (pp. 181–231). Elsevier Academic Press.
- Jost, J. T., & Sterling, J. (2020). The language of politics: Ideological differences in congressional communication on social media and the floor of Congress. *Social Influence*, 15(2), 80–103. <https://doi.org/10.1080/15534510.2020.1871403>
- Kaati, L., Shrestha, A., & Cohen, K. (2016a). *Linguistic analysis of lone offender manifestos* [Paper presentation]. 2016 IEEE International Conference on Cybercrime and Computer Forensic (ICCCF), Vancouver, BC, Canada. <https://doi.org/10.1109/ICCCF.2016.7740427>.
- Kaati, L., Shrestha, A., Cohen, K., & Lindquist, S. (2016b). *Automatic detection of xenophobic narratives: A case study on Swedish alternative media* [Paper presentation]. 2016 IEEE Conference on Intelligence and Security Informatics (ISI), Tucson, Arizona, USA. <https://doi.org/10.1109/ISI.2016.7745454>.
- Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33(2), 125–143. <https://doi.org/10.1177/0261927X13502654>
- Karasawa, M., & Maass, A. (2008). The role of language in the perception of persons and groups. In R. Sorrentino & S. Yamaguchi (Eds.), *Handbook of motivation and cognition across cultures* (pp. 315–341). Elsevier. <https://doi.org/10.1016/B978-0-12-373694-9.00014-3>.
- Keblusek, L., Giles, H., & Maass, A. (2017). Communication and group life: How language and symbols shape intergroup relations. *Group Processes & Intergroup Relations*, 20(5), 632–643. <https://doi.org/10.1177/1368430217708864>
- Kellstedt, P. M. (2003). *The mass media and the dynamics of American racial attitudes*. Cambridge University Press. <https://doi.org/10.1080/10584600500477153>.
- Kocijan, B., & Kuvec, M. (2022). Immigration attitudes among Western and Eastern European MPs: Social identity, economic aspects and political ideology. *Comparative European Politics*, 20, 33–52. <https://doi.org/10.1057/s41295-021-00254-5>
- Körner, R., Overbeck, J. R., Körner, E., & Schütz, A. (2022). How the linguistic styles of Donald Trump and Joe Biden reflect different forms of power. *Journal of Language and Social Psychology*, 41(6), 631–658. [https://doi.org/10.1177/0261927\(22\)1085309](https://doi.org/10.1177/0261927(22)1085309)
- Kriesi, H., Grande, E., Lachat, R., Dolezal, M., Bornschier, S., & Frey, T. (2008). *West European politics in the age of globalization*. Cambridge University Press.

- Lakoff, G. (2004). *Don't think of an elephant! Know your values and frame the debate*. Chelsea Green Publishing Company.
- Larose, D. T. (2014). *Discovering knowledge in data: An introduction to data mining*. Wiley.
- Lee, T. L., & Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the stereotype content model. *International Journal of Intercultural Relations*, 30(6), 751–768. <https://doi.org/10.1016/j.ijintrel.2006.06.005>
- Liu, A. (2022). Pronoun usage as a measure of power personalization: A general theory with evidence from the Chinese-speaking world. *British Journal of Political Science*, 52(3), 1258–1275. <https://doi.org/https://doi.org/10.1017/S0007123421000181>
- Markaki, Y., & Longhi, S. (2013). What determines attitudes to immigration in European countries? An analysis at the regional level. *Migration Studies*, 1(3), 311–337. <https://doi.org/10.1093/migration/mnt015>
- Mendelsohn, J., Budak, C., & Jurgens, D. (2021). Modeling framing in immigration discourse on social media. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2219–2263. <https://aclanthology.org/2021.naacl-main.179>.
- Mendez, M., Cebolla, H., & Pinyol i Jiménez, G. (2014). ¿Han cambiado las percepciones sobre la inmigración en España? *Crítica*, 994, 14–21. <https://fundacionalternativas.org/publicaciones/han-cambiado-las-percepciones-sobre-la-inmigracion-en-espana-2/>
- Meyer, S., & Rosenberger, S. (2015). The politicization of immigration in Austria. In W. Van der Brug, G. D'Amato, D. Ruedin & J. Berkhout (Eds.), *The politicization of migration* (pp. 31–51). Routledge. <https://doi.org/10.4324/9781315723303>
- Morales, L., Pardos-Prado, S., & Ros, V. (2015). Issue emergence and the dynamics of electoral competition around immigration in Spain. *Acta Politica*, 50(4), 461–485. <https://doi.org/10.1057/ap.2014.33>
- Mudde, C. (2007). *Populist radical right parties in Europe*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511492037>.
- Odmalm, P. (2012). Party competition and positions on immigration: Strategic advantages and spatial locations. *Comparative European Politics*, 10(1), 1–22. <https://doi.org/10.1057/cep.2010.20>
- Odmalm, P. (2018). The party politics of migration. In A. Weinar, S. Bonjour & L. Zhyznomirska (Eds.), *The Routledge handbook of the politics of migration in Europe* (pp. 263–295). Routledge. <https://doi.org/10.4324/9781315512853>
- Odmalm, P., & Bale, T. (2015). Immigration into the mainstream: Conflicting ideological streams, strategic reasoning and party competition. *Acta Politica*, 50, 365–378. <https://doi.org/10.1057/ap.2014.28>
- Pardos-Prado, S. (2015). How can mainstream parties prevent niche party success? Center-right parties and the immigration issue. *The Journal of Politics*, 77(2), 352–367. <https://doi.org/10.1086/680040>
- Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. Bloomsbury Press. <https://doi.org/10.1093/lc/fqt006>
- Pennebaker, J. W., & Chung, C. K. (2008). Computerized text analysis of Al-Qaeda transcripts. In K. Krippendorff & M. Bock (Eds.), *A content analysis reader* (pp. 453–465). Sage.
- Pennebaker, J. W., & Chung, C. K. (2012). *Language and social dynamics*. University of Texas Press.
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PLoS ONE*, 9(12), e115844. <https://doi.org/10.1371/journal.pone.0115844>

- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. (2007). *The development and psychometric properties of LIWC2007*. University of Texas.
- Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology*, 59(3), 475–486. <https://doi.org/10.1037/0022-3514.59.3.475>
- Prins, S. B. (2018). *Polarization: identifying extreme attitudes* [Master thesis, University of Twente]. University of Twente Research Repository. <https://essay.utwente.nl/76420/>.
- Ramírez-Esparza, N., Pennebaker, J. W., García, A., & Suriá, R. (2007). Psicología del uso de las palabras: Un programa de computadora que analiza textos en español. *Revista Mexicana de Psicología*, 24(1), 85–99.
- Rinken, S. (2015). Actitudes hacia la inmigración y los inmigrantes: ¿en qué es España excepcional? *Migraciones. Revista del Instituto Universitario de Estudios sobre Migraciones*, 37, 53–74. <https://doi.org/10.14422/mig.i37.y2015.003>
- Ros, V., & Morales, L. (2015). The politicization of immigration in Spain. In W. van der Broug, G. D'Amato, J. Berkhout & D. Ruedin (Eds.), *The politicization of immigration* (pp. 119–139). Routledge.
- Rubini, M., Menegatti, M., & Moscatelli, S. (2014). The strategic role of language abstraction in achieving symbolic and practical goals. *European Review of Social Psychology*, 25(1), 263–313. <https://doi.org/10.1080/10463283.2014.985501>
- Sánchez-Junquera, J., Chulvi, B., Rosso, P., & Ponzetto, S. P. (2021). How do you speak about immigrants? Taxonomy and StereoImmigrants dataset for identifying stereotypes about immigrants. *Applied Sciences*, 11(8), 3610. <https://doi.org/10.3390/app11083610>
- Savaş, Ö., Greenwood, R. M., Blankenship, B. T., Stewart, A. J., & Deaux, K. (2021). All immigrants are not alike: Intersectionality matters in views of immigrant groups. *Journal of Social and Political Psychology*, 9(1), 86–104. <https://doi.org/10.5964/jssp.5575>
- Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(4), 147–141. <https://doi.org/10.1027/1614-2241/a000016>
- Seyle, D. C., & Pennebaker, J. W. (2004). *"We're right, they aren't": A linguistic picture of extremism* [Paper presentation]. Society for Personality and Social Psychology, Austin, TX, USA.
- Shrestha, A., Kaati, L., & Cohen, K. (2020). Extreme adopters in digital communities. *Journal of Threat Assessment and Management*, 7(12), 72–84. <https://doi.org/10.1037/tam0000143>
- Simmons, K., Silver, L., Johnson, C., Taylor, K., & Wike, R. (2018). *In Western Europe, populist parties tap anti-establishment frustration but have little appeal across ideological divide*. Pew Research Center. https://www.pewresearch.org/global/wp-content/uploads/sites/2/2018/07/Pew-Research-Center_Western-Europe-Political-Ideology-Report_2018-07-12.pdf.
- Srinivasa-Desikan, B. (2018). *Natural language processing and computational linguistics: A practical guide to text analysis with Python, Gensim, spaCy, and Keras* (1st ed.). Packt Publishing.
- Sterling, J., Jost, J. T., & Bonneau, R. (2020). Political psycholinguistics: A comprehensive analysis of the language habits of liberal and conservative social media users. *Journal of Personality and Social Psychology*, 118(4), 805–834. <https://doi.org/10.1037/pspp0000275>
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>

- Torregrosa, J., Thorburn, J., Lara-Cabrera, R., Camacho, D., & Trujillo, H. M. (2020). Linguistic analysis of pro-ISIS users on Twitter. *Behavioral Sciences of Terrorism and Political Aggression*, 12(3), 171–185. <https://doi.org/10.1080/19434472.2019.1651751>
- Turnbull-Dugarte, S. J. (2019). Explaining the end of Spanish exceptionalism and electoral support for Vox. *Research & Politics*, 6(2) April–June. <https://doi.org/10.1177/2053168019851680>
- Van der Brug, W., D'Amato, G., Ruedin, D., & Berkhout, J. (Eds.). (2015). *The politicization of migration*. Routledge. <https://doi.org/10.4324/9781315723303>
- Van Dijk, T. A. (1999). Political discourse and racism: Describing others in western parliaments. In S. Riggins (Ed.), *The language and politics of exclusion: Others in discourse* (pp. 31–64). Sage.
- Van Swol, L. M., & Carlson, C. L. (2015). Language use and influence among minority, majority, and homogeneous group members. *Communication Research*, 44(4), 512–529. <https://doi.org/10.1177/0093650215570658>
- Van Swol, L. M., & Kane, A. A. (2019). Language and group processes: An integrative, interdisciplinary review. *Small Group Research*, 50(1), 3–38. <https://doi.org/10.1177/1046496418785019>
- Wei, K., Jacobson López, D., & Wu, S. (2019). The role of language in anti-immigrant prejudice: What can we learn from immigrants' historical experiences? *Social Sciences*, 8(3), 93. <https://doi.org/10.3390/socsci8030093>
- Wissler, L., Almashraee, M., Monett, D., & Paschke, A. (2014). *The gold standard in corpus annotation* [Paper presentation]. 5th IEEE GSC, Passau, Germany. <https://doi.org/10.13140/2.1.4316.3523>.

Author Biographies

Berta Chulvi is a Lecturer in social psychology at the Department of Social Psychology, Universitat de València, Spain. She also holds a PostDoc research position at PRHLT Research Center at Universitat Politècnica de València. Her research interests lie in the connection between computational linguistics and social psychology to enhance the comprehension of intergroup processes.

Mariangeles Molpeceres is an Associate Professor in social psychology at the Department of Social Psychology, Universitat de València, Spain. Her research interests include social representations of minority groups, prejudice and intergroup processes.

María F. Rodrigo is an Associate Professor of the Department of Behavioral Sciences Methodology, Universitat de València, Spain. Her research interests include generalized linear models and applied statistics in different fields such as group decision making, positive youth development and unfounded beliefs.

Alejandro H. Toselli, Ph.D. in Computer Science is currently working as a PostDoc at PRHLT Research Center at Universitat Politècnica de València. His research expertise focuses primarily in the area of Document Analysis and Recognition.

Paolo Rosso, Full Professor in Department of Computer Systems and Computation at the Universitat Politècnica de València. He leads the Natural Language Processing Lab at PRHLT Research Center. His research interests include the automatic detection of harmful information in social media and the spreading of prejudices and stereotypes.