



Selected Papers of #AoIR2024:  
The 25th Annual Conference of the  
Association of Internet Researchers  
Sheffield, UK / 30 Oct - 2 Nov 2024

## TRACING THE SOCIOLINGUISTIC PATTERNS OF POLARIZATION IN THE FACEBOOK DEBATE ON CLIMATE ACTION

Luigi Arminio  
IT University of Copenhagen

Luca Rossi  
IT University of Copenhagen

### Extended Abstract

Sociolinguistic literature has long examined the influence of socialization on communicative norms and practices, considering both semantic and morphosyntactic aspects of communication. Notably, Basil Bernstein's research delineated linguistic production not merely as a byproduct of individual cognitive mechanisms but also as a reflection of the underlying social structures that influence human agency (Bernstein, 1962). In particular, Bernstein (1962) suggested a connection between social stratification and language production, outlining two different sociolinguistic codes: a restricted code, emphasizing informal language and contextual communication, and an elaborated code, characterized by formal language and complex structures, with access determined by social class. However, the influence of socialization on language production may not necessarily be restricted to socioeconomic variables. From this angle, additional theoretical literature suggests the potential connection between political socialization and the communicational patterns of the individuals, theorizing a relationship between the adherence to a certain set of features of a given communicational code (intended as a multidimensional system including linguistic and extralinguistic aspects) and the obtained political success (Ufaeva, 2013). In connection with this, it is notable that the diversity of communication styles in political environments has been empirically studied, with studies drawing on political communication sources such as political party manifestos (Gyasi, 2023) and interviews with politicians (Kayam, 2023). The findings indicated both the restricted accessibility of certain parties' communicational styles (Gyasi, 2023) and the enhanced accessibility of language codes employed by specific political figures (Kayam, 2023). Reflecting on the previously posited and discussed relevance of communicational patterns in political communication, it must be noted that while some studies discuss language codes' impact on political discourse, we can spot a lack of systematic examination of the full range of linguistic features determining the stylistic diversity of social actors and how they impact political polarization dynamics. Addressing this gap facilitates the assessment of the potential impact of sociolinguistic patterns on polarization in public

Suggested Citation (APA): Arminio, L., Rossi, L. (2024, October). Tracing the Sociolinguistic Patterns of Polarization in the Facebook Debate on Climate Action. Paper presented at AoIR2024: The 25th Annual Conference of the Association of Internet Researchers. Sheffield, UK: AoIR. Retrieved from <http://spir.aoir.org>.

debates and may support the development of communication and divulgation approaches that facilitate constructive engagement and bridge ideological divides. Building upon this premise, we tried to measure the sociolinguistic patterns prevalent among stakeholders in the highly polarized climate change debate on social media and their audience. To do that, we compiled a list of active English-speaking actors in the polarized climate debate on social media, categorizing them into pro-climate action (pro-actors) and anti-climate action agents (counter-actors). We constructed the list by drawing from multiple sources including the Desmog climate disinformation database, scholarly research on climate-related social media interactions (Chen et al., 2023), and additional compilations from NPOs, policy insight networks, and content aggregation platforms. Only those actors who displayed recent Facebook activity, defined as having posted at least once in 2023 or 2024, were catalogued. After the list compilation, we resorted to the Crowdtangle API to collect textual content from these actors, capturing posts from March 1, 2024, to February 12, 2024, with a maximum of 100 posts per actor. Overall, we collected data from 10,650 Facebook posts across 250 actors, including 96 pro-actors and 156 counter-actors. Following the data collection, we filtered for climate change relevance with a dictionary-based approach: we compiled climate-related terms from different web sources (including glossaries from institutional and news-related sources), retaining only posts containing at least one of these terms. This yielded a dataset of 2326 posts from 188 actors, comprising 114 counter-actors and 74 pro-actors. Then, posts were labelled based on a comprehensive spectrum of language code indicators, including a readability score, evaluating post ease of comprehension structurally and morphosyntactically (Zamanian, Heydari, 2012); a concreteness score, measuring conceptual depth through a lexicon-based method from a peer-reviewed dictionary (Brysbaert, et. al, 2014); a scientificity score, measuring scientific language utilization using a lexicon-based approach (Suzen et. al, 2019); and a subjectivity score (Verma, 2022), measuring the prevalence of perspectival influence over objective neutrality in each post. In addition to these metrics, we computed social response measures for each post, including the engagement score (Nugroho, Agustina, 2020) gauging audience interest; the emotional polarization score (Muraoka et al., 2021) indicating the direction of the audience's emotional response; and the emotional diversity score (Freeman et al., 2020) signalling the variety of the emotional reactions. Given our data, we aimed to initially examine if there are significant discrepancies in the communicational styles of the two opposing sides of the polarised debate (RQ1) and then, which language code parameters were affecting the audience response to the public communication of both groups of actors (RQ2). To answer the first research question, we measured the point biserial correlation between the position of each climate debate actor (pro vs counter) and each of the measured language parameters. As for RQ2, we computed, for each social response measure, two regression models (one performed on the subset of Facebook data produced by the pro-actors, and one on the analogue counter-actors subset of posts). As for the correlations (Fig. 1), the results showed a moderate correlation ( $r = 0.29$ ) between the climate stances of the actors and the frequency of scientific terminology in their posts, and a weaker association for readability, concreteness, and subjectivity. The direction of the correlation was in favour of the counter-actors only in the case of the subjectivity score. However, the most pronounced relationship is the positive one between the usage of scientific terminology and a stance favouring climate action. Concerning the regression models (Fig. 2),

significant differences in the impact of the language code on the audience response have been detected in the two groups: while in the case of the pro-climate actors, the audience reactions seem to be affected by the communicational style of the posts, the same effect is notably dampened or other times completely absent in the case of the counter actors. These results highlight that while the language register of the actors in the polarized climate debate can affect the receptivity and the emotional response of the digital community, different sides of the polarization spectrum may experience varied impacts, suggesting the necessity for targeted communicational strategies to bridge divides and foster constructive dialogue concerning the climate crisis.

## References

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, 46, 904-911.

Bernstein, B. (1962). Linguistic codes, hesitation phenomena and intelligence. *Language and speech*, 5(1), 31-48. Chen, K., Molder, A. L., Duan, Z., Boulianne, S., Eckart, C.,

Mallari, P., & Yang, D. (2023). How climate movement actors and news media frame climate change and strike: evidence from analyzing twitter and news media discourse from 2018 to 2021. *The International Journal of Press/Politics*, 28(2), 384-413.

Freeman, C., Alhoori, H., & Shahzad, M. (2020). Measuring the diversity of Facebook reactions to research. *Proceedings of the ACM on Human-Computer Interaction*, 4(GROUP), 1-17.

Gyasi, W. K. (2023). The Readability of Political Party Manifestos of the 2016 General Elections in Ghana. *Athens Journal of Mass Media & Communications*, 9(1).

Kayam, O. (2018). The readability and simplicity of Donald Trump's language. *Political Studies Review*, 16(1), 73-88.

Muraoka, T., Montgomery, J., Lucas, C., & Tavits, M. (2021). Love and anger in global party politics: Facebook reactions to political party posts in 79 democracies. *Journal of Quantitative Description: Digital Media*, 1.

Nugroho, A., & Agustina, A. (2020). Examining Corporate Engagement in Social Media: Advancing The Use of Facebook for Corporation Page. *CoverAge: Journal of Strategic Communication*, 10(2), 1-10.

Suzen, N., Mirkes, E. M., & Gorban, A. N. (2019). LScDC-new large scientific dictionary. arXiv preprint arXiv:1912.06858.

Ufaeva, V. (2013). Linguo-Communicative Code of Modern Political Interaction. Historical, Philosophical, Political and Law Sciences, Culturology and Study of Art. Issues of Theory and Practice, 6(32), 192-196.

Verma, P. (2022). Sentiment Analysis Using Subjectivity and Sentiment Score. The Next Normal: Building Agile, Sustainable, Tech-enabled Organizations (TNN-BASTO-2022), 55.

Zamanian, M., & Heydari, P. (2012). Readability of Texts: State of the Art. \*Theory & Practice in Language Studies, 2\*(1).