

War on words: a corpus-based analysis of Indian political neologisms

Apurbalal Senapati¹

Abstract

In the realm of political war (election), language is the most powerful weapon. Political leaders and parties craft and deploy neologisms during their political campaigning and slogans. They always tried to use catchy terms to address a variety of socio-political issues. Newly coined terms or phrases are used in order to address contemporary issues, shape public opinion, rally supporters, and undermine opponents. This *war on words* is particularly intense during election campaigns, where every phrase can sway voters. This study aims to analyze *political neologisms* within the framework of Indian elections using a corpus-based approach for Bengali language. The study is conducted in three phases. The first phase focuses on corpus creation, which includes a political news corpus along with social media texts and comments. The second phase is used to identify the relevant political neologisms. The system utilized an *n-gram model* (for $n=1, 2, 3, 4$) to detect political neologisms. Finally, the identified neologisms are analyzed from a socio-linguistic perspective and documented a total of 327 new political neologisms. Additionally, a detailed description of the twenty most prominent and critically selected neologisms are presented.

Keywords

neologism, linguistics, news corpus, Bengali language, social media text

¹ Central Institute of Technology Kokrajhar, Department of Computer Science and Engineering, India. ORCID: 0000-0001-9124-2563, email: a.senapati@cit.ac.in

Introduction

India is the largest democratic country in the world, with a population of more than 1.4 billion (Sahoo, 2024). The Election Commission of India (ECI) conducts elections at various levels—Parliament (national), Assembly (state), and various regional levels. The Parliamentary elections directly constitute the Lok Sabha, which holds the exclusive authority to elect the Prime Minister and the power to amend or pass new bills (Madhavan, 2024). On the other hand, Assembly Elections select Members of the Legislative Assembly, who represent the state's citizens. They form the State Government, elect the Chief Minister, and hold the authority to oversee law and order at the state level (Ganguly, 2024). India comprises 28 states and 8 union territories, totaling 36 administrative divisions¹. As a result, elections take place throughout the year in the country. In the last 2024 Parliament election, 65.79% voter turnout was recorded at polling (GE, 2024).

Politics plays a crucial role in contemporary democracy, serving as the cornerstone that shapes societal functioning. It's the mechanism through which power is allocated, recognizing and influencing social advancements. Socio-political concerns such as educational and healthcare access are products of both societal disparities and political choices. Analyzing these matters through a socio-political perspective allows us to grasp the political underpinnings of societal challenges and the social impacts of political remedies. In India, elections are influenced not only by political and economic issues but also by factors such as religion, caste, and regional identities. These factors often play a significant role in shaping voter behavior and political strategies. Political parties frequently appeal to caste-based vote banks or religious sentiments to garner support, a practice that has been widely documented in academic research and political analyses (Patil, 2022).

In the diverse landscape of Indian politics, use of language act as both a weapon and a shield. To captivate the public mind political parties, politicians, and public figures craft and manipulate words to frame narratives in favors of their philosophy or ideology. In this context, this paper aims to explore the phenomenon of political neologisms coined during election campaigns. A corpus-based experiment has been conducted to analyze these newly coined terms and their impact.

1. Social media and political neologism

In recent times, a significant portion of the population has been active on various social media platforms, utilizing them as avenues to voice their desires, protest, and campaign². The electoral landscape and political party propaganda has also undergone substantial changes over the past two decades. Consequently, numerous new terminologies have arisen to meet the demands of the evolving socio-political framework. In some instances, political leaders have demonstrated intolerance by indirectly using unethical or unconstitutional discourse (Carme et al., 2021), thereby giving rise to new neologisms. This paper delves into the realm of political neologisms. To achieve this, we begin by constructing a political news corpus along with social media text and employing a method to identify, interpret, and analyze new terminologies and their usage. It also highlights the impact of socio-political culture and examines how cultural factors aid politicians in the creation of political neologisms.

¹ <https://knowindia.india.gov.in/states-uts/> (Accessed 15.03.2025).

² <https://datareportal.com/reports/digital-2024-global-overview-report> (Accessed 25.01.2025).

Political neologisms can originate from various sources such as political leaders, social media wings of the political party, activists, media houses, or the general public and can quickly spread through social media, political rally or other communication channels. The primary concern of political neologisms is their capacity to shape public opinion and influence political narratives (Zailani, 2019, Saja et al., 2019). They can serve as rallying cries for movements, encapsulate policy agendas, or label opponents in a particular way. Sometimes these neologisms are used as a political tool to spread their opinions, ideology, or the sentiment of a specific community.

With social media's rise, political parties have embraced dedicated social media teams for campaigning (Rodrigues, 2017). These platforms have become vital tools for reaching and interacting with voters. They offer unique benefits and challenges, integrating seamlessly into modern political communication strategies. As the political environment shifts, new terms emerge to address and counter opposing viewpoints (Rodrigues, 2020, Chadwick, 2013). Similar to any other domain of linguistics, the language of politics undergoes continual evolution. These neologisms wield significant influence, molding our comprehension and involvement in political matters.

2. State of the art in neologism detection

Linguistics has been addressing the issue of neologism for a long time, initially relying on manual methods. Lexicographers manually analyzed texts from diverse sources such as literature, newspapers, technical documents, and scientific texts (Frank, 1991). The advent of digital technology has revolutionized linguistic and natural language processing (NLP) research, leading to the creation of novel data collection methods and tools (Mejri et al., 2011; Humbley et al., 2016). Nowadays, a variety of automated techniques and tools are accessible for scanning vast amounts of text, particularly from newspapers and blogs, to automatically identify newly coined words in various languages (Ted, 1993; Paul et al., 2000; Ingrid et al., 2014; Maarten 2022). This section provides a brief overview of some of the significant advancements in this specific field.

Corpora have frequently used resources in the study of neologisms (Laurie et al., 2000; Hohenhaus, 2006), with researchers often employing various methods such as the web-as-corpus approach, news corpora (Renouf, 2007), or resources like Wikipedia (Tony et al., 2010). While cloud sources provide a substantial amount of text for studying neologisms, there are various challenges associated with it (Senapati et al., 2022). The text scraped from web sources includes various types of noise such as image information, links, informal language, unknown characters, numeric data, misspellings, etc.

Paryzek (2008) explores different extraction methods for identifying neologisms, while Jack et al. (Jack et al., 2016) discuss the factors influencing the rise and acceptance of new words. Daphne et al. (Kerremans et al., 2012) integrated corpora from social media platforms into their research. Some organizations (Chen et al., 2020) compiling scientific COVID-19 literature are referred to as corpora. Alyeksyeyeva et al. (2020) conducted a study on linguistic aspects of the English language, identifying various jargon, slang, and terms related to COVID-19 that have become culturally adopted. Wang (Wang et al., 2024) added 1,204 neologisms to the Oxford Advanced Learner's Dictionary. They conducted part-of-speech analysis and classified these neologisms according to their word formation. The results showed that 71% of them were nouns. Zhou (Zhou et al., 2023) compiled coronavirus-related neologisms and addressed lexicographic issues in thirteen research papers in various languages, such as English, Hungarian, Korean, Italian, German, etc. The study also emphasized the cross-cultural

perspectives of lexicographers from diverse linguistic backgrounds. Hilal (Hilal, 2019) focuses on French neologism in the context of Arabic–French general bilingual dictionary. Senapati (Senapati, 2023) conducted a semi-automated corpus-based study on COVID-19 neologisms. Initially, they created a web-based news corpus and then employed dictionary entry strategies to identify neologisms. Zailani (2019) researched political neologisms from a sociolinguistic perspective. Estabraq (2020) focused on neologisms in social media, analyzing three million tweets. Kernerman (Klosa et al., 2020) edited a special issue of dictionaries dealing with neological lexicography in various languages like Danish, Dutch, Estonian, Frisian, Greek, Korean, Spanish, and Swahili. It addresses relevant aspects of lexicography in those languages, presenting state-of-the-art research into neology and ideas about modern lexicographic treatment of neologisms in various dictionary types. Certain researchers are endeavoring to tackle the translation challenges associated with political neologisms (Hanaqtah, 2019). Mestre endeavored to pinpoint the emotional lexicon linked to the political conflict within the Ukrainian war (Mestre-Mestre, 2023). Sharhan (Sharhan, 2022) carries out a political discourse analysis of the special political circumstances of Iraq. The literature review uncovered a scarcity of work on neologisms in South Asian languages, particularly on political neologisms, which are rarely addressed. This situation motivated us to explore political neologisms within the political discourse of India, the world’s largest democratic country (Verma, 2024).

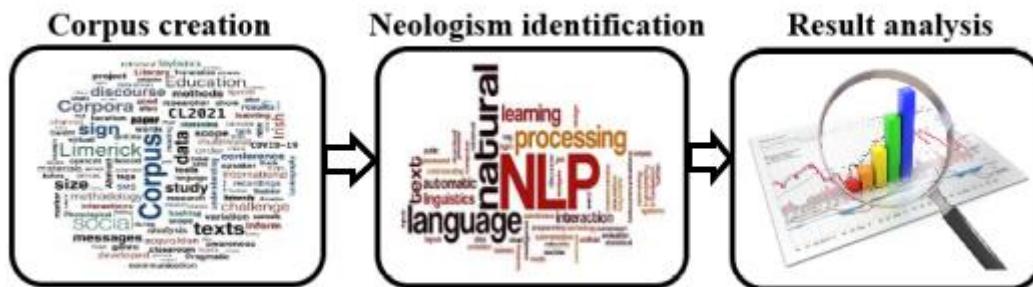
This article presents an approach that significantly diverges from traditional methods of neologism detection. It focuses specifically on identifying political neologisms within news and social media contexts. Such neologisms tend to proliferate during state or national elections due to the prevailing socio-political-cultural climate. However, gathering a dedicated corpus from social media texts poses a significant challenge. Unlike newspapers, where articles are authored by third parties, social media content is user-generated, introducing potential biases and information gaps. Despite these obstacles, our current efforts prioritize detecting neologisms within news and social-media texts using the *n-gram model*. To summarize, this paper’s contribution can be outlined as follows:

- Compile an exclusive political news corpus from diverse news sources and social media commentary,
- Utilized an n-gram model to detect political neologisms,
- Identified and documented 327 political neologisms,
- Conducted a study to explore the differences between political neologisms and other types of neologisms and elaborated with examples,
- Analyzing the socio-political factors to explain the emergence of such neologisms.

3. System description and methodology

The entire study comprises three distinct steps. The first involves creating a corpus, followed by developing a model to identify neologisms within that corpus and culminates in analyzing the results. Each step is elaborated upon in the following subsections. The system’s block diagram is illustrated in Figure 1.

Figure 1. Block diagram of the system for the political neologism.



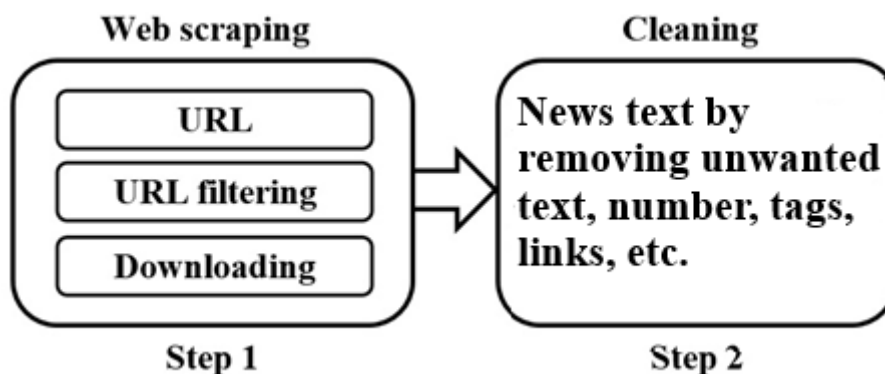
Source: author's own research.

3.1. Data preparation

The experiment is conducted in Bengali (Census, 2021) a major languages spoken in India. To create the political news corpus, we have collected political news articles that cover the election campaigns of different political parties. Throughout the election period, newspapers prioritize covering election-related news extensively. As a result, it is assumed that these news articles contain a wealth of political terminology and jargon. In our experiment, we considered newspapers in Bengali languages and from various social media texts and comments.

In this context, we included political news articles from leading daily newspapers such as *Anandabazar Patrika*³ in Bengali language. The Election Commission of India announced the election date on March 16, 2024 (ECI, 2024), leading to the collection of news corpus data from March 1, 2024, to May 20, 2024, spanning a total of 81 days. Figure 2 shows the block diagram of the system used for political corpus building.

Figure 2. Data preparation: web scraping and cleaning the political news corpus.



Source: author's own research.

Web scraping

- To access political news articles from past dates, we initially compiled a list of Uniform Resource Locators (URL) from the online web portal of the respective newspapers.
- To specifically choose political news articles, we applied a filter to the URLs. This filter is based on a list of keywords name of political parties, political leaders of various parties,

³ <https://www.anandabazar.com/> (Accessed 25.01.2024).

political issues, election-related terminologies, etc. related to the elections. If a URL contains any of these keywords, it remains unchanged; otherwise, it is excluded. After passing the prepared list of URLs through this filter, obtain the final URL list for news articles related to the elections.

- We utilized a Python-based web scraper using the *Beautiful Soup*⁴ library to download the news articles. After successfully employing the web scraper, we saved each article’s raw text as an individual *.txt* file.

Cleaning

- Process each *.txt* file containing raw news text by removing unwanted elements such as HTML tags, numerical values, links, advertisement text, and special characters.
- Once all unwanted elements have been removed, extract the essential details i.e. actual new text. Specifically, it retains only the news headline with its text, saving this information in a *.txt* file using UTF-8 encoding.
- Each file in the corpus will consist of a news article’s headline, and publishing date, followed by its content. Therefore, the total number of files in the corpus is equal to the total number of news articles contained within it.

Corpus summary

As mentioned earlier, the political corpus is generated from the two specified newspapers. Table 1 shows the brief volume of the political news corpus of Bengali. Moreover, we have collected 10,000 lines of user comments and political texts in Bengali. This data is gathered from various social media political groups and discussions on contemporary political issues those are not shown in Table 1.

Table 1. Volume summary of political news corpus.

Sl. No.	Description	Bengali News Corpus
1	Number of news articles	1,500
2	Number of sentences	40,430
3	Number of words	5,49,909
4	Disk space	9.2 MB

Source: author’s own research.

3.2. Methodology

Once the political news corpus has already been prepared, the next move into the neology identification phase (shown in Figure 1). An *n-gram* refers to a consecutive set of *n*-words within a text. This *n-gram* sequence is straightforward and powerful for examining and anticipating word sequences, thereby establishing themselves as fundamental instruments across a range of NLP tasks (Jurafsky et al., 2024). Political neologisms encompass not just

⁴ <https://www.crummy.com/software/BeautifulSoup/bs4/doc/> (Accessed 25.01.2024).

single words but also multiword expressions, making the *n-gram* sequence particularly suitable for handling this type of linguistic phenomenon. Note that the number of *n-gram* in this context is extensive and poses challenges for manual investigation. To streamline this list, entities containing stop words, named entities, and similar elements are eliminated. The Bengali stop-word list is sourced from the FIRE⁵ and Kaggle⁶. This is particularly necessary in political news where named entities like political parties, leaders, locations, etc., appear frequently. In the implementation of *n-gram* sequence, we have used the Python natural language toolkit library (Nellis et al., 2016).

Preparing n-gram sequence

- To identify the neologism, a semi-automated technique is used where first *n-gram* (for $n=1, 2, 3, 4$) are created from the entire corpus.
- Eliminate all *n-grams* that include stop words, named entities or punctuation marks; hence it substantially reduces the *n-gram* list.

Automated approach for neology detection

Philosophically, newly coined terms should not yet have entries in dictionaries of contemporary times. This principle has been leveraged to identify neologisms from a corpus through automated methods. Check the online dictionary entries for the *n-grams*; if none are found, it could be a neologism, which is then be verified manually. An online Bengali-English dictionary⁷ was used to verify the dictionary entries. This method is noted to identify only pure neologisms but fails to detect existing words with new meanings. Such cases are identified by manually analyzing their context.

4. Result and discussion

A total of 327 distinct political neologisms have been identified, with a selection of the most prominent ones displayed in Table 2. Out of the 327 neologisms, twenty have been meticulously selected and are presented in Table 2. These neologisms encompass a wide range of categories and were not chosen randomly. It's worth noting that many of these neologisms utilize existing terminologies (SI No. 2, 11, 12, 13 and 19 in Table 2) but with altered interpretations within the realm of political discourse. It is also noteworthy that many of these neologisms are employed to convey sarcasm, spread hateful intentions, or make promises. Below are the detailed analyses of these neologisms. It has been noted that out of the 327 neologisms identified, inflectional forms are not included.

Among these, 23 are classified as pure neologisms, i.e. they are newly coined. While the remaining fall into other categories such as borrowed words, derived terms, blends, and so on. Additionally, it is noted that the majority of these pure neologisms are found in social media texts. Handling social media text is more complex. The key challenges include managing various content types such as videos, images, emojis, links, and hashtags. Filtering out noise and irrelevant information like slang, spam, misspellings, and abbreviations; dealing with

⁵ <https://www.isical.ac.in/%E2%88%BCfire/data/stopwords> (Accessed 25.01.2025).

⁶ <https://www.kaggle.com/datasets/towhidahmedfoysal/bangla-stop-word-list> (Accessed 25.01.2025).

⁷ <https://www.english-bangla.com/bntoen> (Accessed 24.05.2024).

ambiguity and context-dependency in text. Another issue is the code-mixing, where multiple languages are used, often requiring manual verification of these texts.

Note that, while an n-gram may be identified as a potential neologism from this list, it does not necessarily mean that it will be confirmed as such. To ensure confirmation, its context is analyzed, and based on this analysis, it is either accepted or rejected. This process involves retrieving the text containing the neologism and examining its context. Some entries in Table 2 (1, 3, 10, 16 and 20) have a blank entry under the column "Literary meaning". This indicates that these are purely newly coined terms with no other established literary meaning as of now in Bengali; they exist solely with their neological meaning given in the next column.

Table 2. Selected political neologisms of various categories.

Sl. No.	Political neologism	Literary meaning	Interpretation in political discourse
1	সেকু / Seku	Distortion form of secularism
2	ভাইপো / Vaipo	Nephew	Indication of a particular political leader in a Sarcastic way
3	গোদি মিডিয়া / Godi media	Paid media / biased media blindly support to the ruling party
4	মিত্রোঁ / Mitron	Friend	Dear countryman / citizen
5	গাদ্দার / Gaddar	Traitor	A leader who has switched their political party
6	৫৬ ইঞ্চি / 56 inch	56-inch length	A particular Prime Minister of India
7	রামলালা / Ramlala	Lord Ram	A particular idol of Lord Ram in a specific temple that was recently inaugurated at Ajodhya
8	হার্মাদ / Harmad	Goons	Cadre of a specific political group
9	গোসন্তান / Gosantan	Calf	Political or religious person who do not utilize their brain
10	বিজেমুল / Bijemool	Blended political activity and ideology
11	দুধেল গাই / Dudhel gai	Milky cow	Vote bank of a specific group
12	খেলা হবে / Khela hobe	Will play	Unscrupulous tactic to gain an advantage in the game
13	জয় শ্রী রাম / Jai Sri Ram	A religious slogan / Lord Ram	Used for teasing to the different ideology
14	হোয়াটসঅ্যাপ ইউনিভার্সিটি / WhatsApp University	Learning from WhatsApp	Rely on unverified or false information circulated in social media / spread of misinformation, rumors, or unverified information through messaging platforms like WhatsApp, Facebook, based on blind faith.
15	আবকি বার ৪০০ পার / Aabki bar 400 par	This time 400 crossed	Securing over 400 seats in the upcoming election
16	ঘুষপেটিয়া / Ghuspetia	Refugee / Illegal infiltrator or intruder. Generally used in the context of illegal immigration and infiltration in border areas
17	অতিবাম / Atibam	ultra-left	Leftist parties and leftist student fronts across various universities
18	চাড়ি / Chaddi	Underwear	Generally used sarcastically, ironically, or contemptuously to refer to supporters of a specific Hindutva or right-wing organizations
19	রোহিঙ্গা / Rohingya	An ethnic group of Myanmar	A vote bank consisting of a specific minority refugee group
20	ফেকু / Feku	Distortion form of fake / Used sarcastically to target a specific political leader

Source: author's own research.

4.1. Distinctive characteristics of political neologisms

Our research has identified key distinctions between political neologisms and their conventional counterparts.

An important aspect of political neologisms is that they are used as-is across languages, without translation. The neologisms listed in Table 2 are the same for both Bengali and Hindi languages (it is verified in the political Hindi articles in *Jagran*⁸, a daily Hindi newspaper during the election campaigning). For example, the neologisms listed in Table 2 are employed consistently across multiple languages.

Political neologisms have a short lifespan. In most cases, a variety of neologisms are coined based on the socio-political situation of a particular election. Once the election is over, many of these neologisms become irrelevant and naturally become obsolete. As an illustration, the neologism *kala dhon*, which literally means black money but has a neological meaning referring to *Indian black money deposited in Swiss bank*, was widely used during the 2014 election campaign. However, after the election, it has become rarely used and consequently obsolete.

The majority of political neologisms are constructed using existing terminologies, with only a few of them being genuinely newly coined. For example, in the aforementioned, entry 10 in Table 2, the neologism *Bijemool* is newly coined, and therefore, it does not have any established literary meaning.

Most of the political neologisms employed in context involve hate speech, sarcasm, criticism, or political slogans. For example, among the neologisms listed in Table 2, the first 14 are used in contexts related to hate speech, sarcasm, and criticism, while the 15th one functions as a political slogan.

4.2. Etymological study of political neologism

Political neologisms arise for several reasons, and based on our study and analysis of the results, some of the primary factors can be identified.

One key reason is the emergence of new concepts in response to the evolving political terrain, where novel ideologies, policies, and phenomena demand new terminology. For instance, the term *Bijemool* (entry 10 in Table 2) is a blend of the names of two ideologically distinct political parties—Trinamool and BJP—and is used to describe individuals who align with both parties. Another important factor is the role of slogans and branding in political propaganda and mobilization. In the current parliamentary election in India, parties have created slogans like *Abki Baar 400 Paar* (Gaikwad et al., 2021), which literally means "this time cross 400," referring to the opposition's challenge to the ruling alliance to secure more than 400 of 543 seats. Neologisms also serve to enhance communication and clarity, offering concise ways to express complex issues. An example is *WhatsApp University*, which denotes reliance on unverified information circulated via social media instead of critical thinking. Additionally, community sentiment frequently influences neologism creation, especially in Indian politics (Nellis et al., 2016). The term *Dudhel gai* (entry 11 in Table 2) targets a specific minority community in this context. Similarly, religious sentiment plays a major role in generating terms

⁸ <https://www.jagran.com/> (Accessed 25.01.2024).

like *Jai Sri Ram* and *Gosantan* (Table 2), which resonate with religious identity and mobilization (Gaikwad et al., 2021).

Another mechanism involves the deliberate distortion of existing words to express contradiction or disdain, particularly towards philosophical concepts such as secularism and Marxism. For example, *Seku* and *Maku* (entry 1 in Table 2) are common on social media, derived by altering the original terms. Some neologisms involve borrowing from other languages and assigning new meanings, as exemplified by *Harmad* (entry 8 in Table 2). Political actors also frame debates strategically through coined terms—*Godi media* (entry 3 in Table 2), for example, characterizes media perceived as loyal to the ruling party. Neologisms may also be used for their catchiness and memorability, helping to disseminate political messages and engage supporters; *Mitron* (entry 3) is one such term, aimed at fostering camaraderie and closeness with audiences. In more extreme cases, neologisms express hatred toward refugee groups, as illustrated by entries 16 and 19 in Table 2, including *Rohingya*, which has come to be used politically to suggest "illegal voter" manipulated by opposition parties. Lastly, certain neologisms function as censorship terms, where offensive or sensitive content is partially masked with characters such as asterisks, dollar signs, or dashes. Examples like বোকা*দা / Boka*da and বা*ল / Ba*1 exemplify such obscured expressions. These have yet to be fully recognized or analyzed as neologisms but represent a growing area of interest in political language.

Conclusion

57

This paper offers significant contributions across several dimensions. It not only pioneers the creation of political neologisms but also documents and conducts a sociolinguistic analysis of them. It can be summarized as a dedicated political news corpus has been developed. n-gram sequence is employed to extract political neologisms. The system's results have been meticulously examined and documented, resulting in a list of 327 political neologisms. The neologisms are categorized based on their etymology.

Furthermore, it conducted a detailed analysis of the distinction between political neologisms and conventional neologisms, making a unique contribution in this domain. Additionally, it also focuses on both pure neologisms and hateful neologisms primarily originate from text found on social media platforms. This finding will contribute to policymaking efforts in regulating hate speech and protecting human rights on social media, particularly concerning refugees and minorities.

References

- Alyeksyeyeva, I., Chaiuk, T. A., & Galitska, C. A. (2020). Coronaspeak as key to coronaculture: Studying new cultural practices through neologisms. *International Journal of English Linguistics*, 10(6), 202–212. <https://www.ccsenet.org/journal/index.php/ijel/article/view/0/43724>
- Carme, C., S'anchez-Miret, H., & Youngs, R. (2021). *The impact of disinformation on democratic processes and human rights in the world* (Policy Department for External Relations, Directorate General for External Policies of the Union, PE 653.635). European Parliament. <https://doi.org/10.2861/59161>
- Census of India. (2021). *Census of India 2011: National Population Register & Socio Economic and Caste Census* (ORGI PPT 2011 38).

- Chadwick, A. (2013). *The hybrid media system: Politics and Power* (1st ed.). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199759477.001.0001>
- Chen, L. C., Chang, K. H., & Chung, H. Y. (2020). A Novel Statistic-Based Corpus Machine Processing Approach to Refine a Big Textual Data: An ESP Case of COVID-19 News Reports. *Applied Sciences*, 10(16), 5505. <https://www.mdpi.com/2076-3417/10/16/5505>
- Kerremans, D., Stegmayr, S., & Schmid, H.-J. (2012). The NeoCrawler: Identifying and retrieving neologisms from the internet and monitoring ongoing change. In K. Allan & J. A. Robinson (Eds.), *Current methods in historical semantics* (pp. 59–96). Mouton de Gruyter.
- Donovan, R., & Neill, M. (2008). A systematic approach to the selection of neologisms for inclusion in a large monolingual dictionary. In *Proceedings of the 13th EURALEX International Congress* (pp. 571–579).
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), 61–74.
- Election Commission of India [ECI]. (2024, March 16). *Schedule of elections: General Elections to Lok Sabha* (No. ECI/PN/23/2024). Nirvachan Sadan. <https://elections24.eci.gov.in/docs/press-note-no-23.pdf>
- Estabraq, R. I., Al-Khafaji, S. A., Hussein, H. M., & Al-Haneen, A. H. (2020). A sociolinguistic approach to linguistic changes since the COVID-19 pandemic outbreak. *Multicultural Education*. <https://doi.org/10.5281/zenodo.4262696>
- Frank, A. S. (1991). From n-grams to collocations: An evaluation of Xtract. In *Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics* (pp. 279–284).
- Gaikwad, N., & Nellis, G. (2021). Do politicians discriminate against internal migrants? Evidence from nationwide field experiments in India. *American Journal of Political Science*, 65(4), 790–806.
- Ganguly, S., & Eswaran, S. (Eds.). (2024). *The Oxford handbook of Indian politics* (Online ed., Oxford Academic, 19 September 2024). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198894261.001.0001>
- Government of India, Election Commission [GE]. (2024). *Press note No. ECI/PN/116/2024*. Nirvachan Sadan, Ashoka Road, New Delhi-110001. <https://elections24.eci.gov.in/eci-updates.html>
- Hanaqtah, M. F. (2019). Translation of political neologisms coined by politicians: Issues and strategies. *Journal of Social Sciences (COES&RJ-JSS)*, 8(1), 157–168. <https://doi.org/10.25255/jss.2019.8.1.157.168>
- Hilal, I. (2019). La création du néologisme français dans le dictionnaire général bilingue arabe–français moderne. *Lexikos*, 29(1). <https://doi.org/10.5788/29-1-1523>
- Hohenhaus, P. (2006). Bouncebackability: A web-as-corpus-based study of a new formation, its interpretation, generalization/spread, and subsequent decline. *SKASE Journal of Theoretical Linguistics*, 3, 17–27.
- Humbley, J., & Sablayrolles, J. F. (2016). Néologie et corpus. *Neologica. Revue Internationale de Néologie*, 10.
- Ingrid, F., Delphine, B., & Christophe, G. (2014). From non-word to new word: Automatically identifying neologisms in French newspapers. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (pp. 4337–4344). http://www.lrec-conf.org/proceedings/lrec2014/pdf/288_Paper.pdf
- Grieve, J., Nini, A., & Guo, A. (2017). Analyzing lexical emergence in modern American English online. *English Language and Linguistics*, 21(1), 99–127. <https://doi.org/10.1017/S1360674316000113>
- Jack, G. Andrea, N. and Diansheng, G. (2016). Analyzing lexical emergence in Modern American English online. In: english language and linguistics.
- Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing* (3rd ed., draft). Pearson. <https://web.stanford.edu/~jurafsky/slp3/>
- Kessler, M. (1993). Legal discourse and political intolerance: The ideology of clear and present danger. *Law & Society Review*, 27(3), 559–597. <https://doi.org/10.2307/3054105>
- Klosa-Kückelhaus, A., & Kernerman, I. (2020). Global viewpoints on lexicography and neologisms: An introduction. *Dictionaries: Journal of the Dictionary Society of North America*, 41(1), 1–9. <https://doi.org/10.1353/dic.2020.0001>
- Laurie, B., & Renouf, A. (2000). Contextual clues to word-meaning. *International Journal of Corpus Linguistics*, 5, 231–258.
- Jongejan, M. (2022). NeoTag: A POS tagger for grammatical neologism detection. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 2118–2124). http://www.lrec-conf.org/proceedings/lrec2012/pdf/1098_Paper.pdf
- Maarten, J. (2022). NeoTag: a POS Tagger for Grammatical Neologism Detection. In: Proceedings of the Eighth International Conference on language resources and evaluation (LREC'12), pp. 2118–124, http://www.lrecconf.org/proceedings/lrec2012/pdf/1098_Paper.pdf (accessed 26-01-2025).
- Madhavan, M. R. (2024). The Lok Sabha. In S. Ganguly & E. Sridharan (Eds.), *The Oxford handbook of Indian politics* (online ed.). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198894261.013.3>
- Mejri, S., & Sablayrolles, J. F. (2011). Présentation: Néologie, nouveaux modèles théoriques et NTIC. *Langages*, 183, 3–9.

- Mestre-Mestre, E. M. (2023). Emotion in politics in times of war: A corpus pragmatics study. *Corpus Pragmatics*, 7, 323–344. <https://doi.org/10.1007/s41701-023-00147-w>
- Nellis, G., Weaver, M., & Rosenzweig, S. C. (2016). Do parties matter for ethnic violence? Evidence from India. *Quarterly Journal of Political Science*, 11(3), 249–277. <http://dx.doi.org/10.1561/100.00015051>
- Patil, G. (2022). Caste, religion and language as drivers of voting patterns in Indian politics. *Educational Administration: Theory and Practice*, 28(4), 350–353. <https://doi.org/10.53555/kuvey.v28i4.8020>
- Paul, R., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the Workshop on Computer Corpora* (pp. 1–6).
- Paryzek, P. (2008). Comparison of selected methods for the retrieval of neologisms. *Investigaciones Lingüísticas*, 16, 163–181.
- Renouf, A. (2007). Tracing lexical productivity and creativity in the British media: The Chavs and the Chav-Nots. In J. Munat (Ed.), *Lexical creativity, texts and contexts* (pp. 61–92). John Benjamins. <https://benjamins.com/catalog/sfsl.58.12ren>
- Rodrigues, U. (2020). Political communication on social media platforms. In A. Athique & V. Parthasarathi (Eds.), *Platform capitalism in India* (pp. [xx–xx]). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-44563-8_11
- Rodrigues, U., & Niemann, M. (2017). Social media as a platform for incessant political communication: A case study of Modi’s “Clean India” campaign [Journal contribution]. Deakin University. <https://hdl.handle.net/10536/DRO/DU:30102388>
- Sahoo, N. (2024). Political finance in the digital age: The case of India. *International IDEA*. <https://doi.org/10.31752/idea.2024.30>
- Saja, A., & Hadi, A. (2019). The significance of political neologisms. *Journal of Education College Wasit University*, 2(37), Article 10. <https://doi.org/10.31185/eduj.Vol2.Iss37.1107>
- Senapati, A., & Nag, A. (2022). Neologism related to COVID-19 pandemic: A corpus-based study for the Bengali language. In V. Bhateja, J. Tang, S. C. Satapathy, P. Peer, & R. Das (Eds.), *Evolution in computational intelligence* (Smart Innovation, Systems and Technologies, Vol. 267, pp. [xxx–xxx]). Springer. https://doi.org/10.1007/978-981-16-6616-2_20
- Senapati, A. (2023). A semi-automated approach for Bengali neologism. *SN Computer Science*, 4, 428. <https://doi.org/10.1007/s42979-023-01866-2>
- Sharhan, K. S. (2022). A political discourse analysis of the speeches of Iraqi designated prime ministers. *Language, Discourse & Society*, 10(2), Article 20. <https://doi.org/10.5281/zenodo.10207932>
- Veale, T., & Butnariu, C. (2010). Harvesting and understanding online neologisms. In S. Olsen (Ed.), *Cognitive perspectives on word formation* (pp. 399–418). Mouton de Gruyter.
- Verma, A., Sridhar, S., Dar, K. F., & Asthana, M. K. (2024). Scope and challenges in the implementation of Time Bank in India: A qualitative study. *Discover Global Society*, 2, Article 14. <https://doi.org/10.1007/s44282-024-00040-0>
- Wang, A., & Chen, X. (2024). On the inclusion of neologisms in *Oxford Advanced Learner’s Dictionary* (10th edition). *Lexikos*, 34(1), 41–50. <https://doi.org/10.5788/34-1-1871>
- Zailani, A. A. (2019). Coinage and neologism in Hausa political programs: A sociolinguistics perspective. *Register Journal*, 12(2), 294–311. <https://doi.org/10.18326/rgt.v12i2.294-311>
- Zhou, J., & Dong, S. (2023). Annette Klosa-Kückelhaus and Ilan Kernerman (Eds.), *Lexicography of coronavirus-related neologisms*. *Lexikos*, 33(1), 131–135. <https://doi.org/10.5788/33-1-1801>