ACCSCIENCE PUBLISHING

International Journal of Population Studies

REVIEW ARTICLE

Framing migration in UK and US media: A corpus-based critical discourse analysis

Tongkai Xu* and **Fei Yang**

School of Liberal Arts, Qilu Institute of Technology, Jining, Shandong, China

Abstract

Migration is among the most contested issues in public debate in the United Kingdom (UK) and the United States (US), yet how far its media representation is ideologically structured, and how stable that structure is over time, remain open questions. This study combines corpus linguistics and critical discourse analysis to examine frequency, collocation, evaluative polarity, agency attribution, and metaphor around the lemmas "migrant," "refugee," and "immigrant" in a corpus of 500 news-media articles (approximately 1.2 million words) published between 2015 and 2024 by two UK outlets (*The Guardian* and *The Telegraph*) and two US outlets (*The New York Times* and Fox News). Change over time is analyzed across four periods rather than single years. Two principal findings emerge. First, migration coverage is ideologically patterned: evaluative polarity around migrant differs significantly across outlets ($\chi^2(3) = 54.17$, $p < 0.001$, Cramér's $V = 0.52$), with right-leaning outlets framing migration through labels of "illegality," invasion, and enforcement, and left-leaning outlets through rights, journeys, and structural causes. Much of this evaluative work is done by the choice of label itself: outlet differences were smaller for refugee (Cramér's $V = 0.26$) than for migrant, while immigrant showed a renewed ideological split ($V = 0.40$) structured by the contrast between undocumented and illegal. Second, this patterning is neither stable nor sufficient as an explanation. In event-dominated periods, variation within outlets over time was comparable to the variation between them, and the 2022 Ukrainian refugee crisis produced convergence on humanitarian framing across the ideological spectrum that was not extended to Syrian or Central American refugees. In addition, the frequent use of nominalization and agentless passives in enforcement reporting presents deportation and detention as administrative facts rather than political decisions. Migration framing in UK and US media is therefore best understood as the joint product of outlet ideology, event context, and the racialized and geopolitical positioning of particular migrant populations. The decade-long comparative design suggests that selective humanitarianism, rather than partisan difference alone, deserves to be a central concern of research and professional practice.

Keywords: Critical discourse analysis; Corpus linguistics; Media discourse; Migration; Ideology; Selective humanitarianism; United Kingdom; United States

***Corresponding author:**
Tongkai Xu
(15864907623@163.com)

Citation: Xu, T., & Yang, F. Framing migration in UK and US media: A corpus-based critical discourse analysis. *International Journal of Population Studies*. doi: 10.36922/IJPS026280168

Received: July 8, 2026

Revised: October 21, 2026

Accepted: October 23, 2026

Published online: November 18, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

1.1. Migration, media framing, and ideological representation

Migration is one of the most contested issues in public debate in the United Kingdom (UK) and the United States (US), and the news media play a central role in shaping how questions of border control, asylum adjudication, and integration are framed in both countries. In the UK, net migration has become a key measure of government performance, and successive commitments to reduce it have been the subject of sustained controversy (Financial Times, 2025). In the US, the southern border has for much of the last decade stood in for immigration policy as a whole. In both countries, migration receives extensive media coverage, political agendas are clearly reflected in reporting, and terms such as "infiltrators" and "border crossers" are far from neutral labels for people who cross borders. This study begins from the assumption that the media are not merely products of these debates but participants in them. News media are active agents that choose what to report, who has a voice, and which words name and characterize people on the move (Colbern *et al.*, 2024; Liu, 2021; Phelan, 2022).

The choices between the expressions "undocumented immigrant" and "illegal alien," between "refugee" and "migrant," and between "arriving" and "flooding in" are not merely stylistic. They each make claims of legitimacy and deservingness (Crawley *et al.*, 2016; De Genova, 2013; Kelling & Motschenbacher, 2018), and—as recent publications in this journal have emphasized—these categories are themselves problematic formulations and not neutral descriptions of mobility (Keyel, 2024). A single axis of difference has dominated the literature. Right-leaning media tend to portray migrants as a threat to security, cultural cohesion, and public resources, often through metaphors of invasion and inundation that have a dehumanizing effect (Markowitz & Slovic, 2020; Parker *et al.*, 2022; Taylor, 2021). Left-leaning outlets, by contrast, foreground displacement, suffering, and rights-based claims. Working with 47,388 US articles, Chen *et al.* (2022) found that right-leaning coverage of Syrian refugees was organized around security risk and left-leaning coverage around humanitarian deservingness. Eberl *et al.* (2018) report parallel divergences across European media systems, and Liu (2021) traces comparable patterning in a four-system comparison that includes both the UK and the US. In their foundational analysis of the UK press between 1996 and 2005, Gabrielatos and Baker (2008) showed that these associations were not only present but durable, and that the terms "refugees," "asylum seekers," "immigrants," and "migrants" were used almost interchangeably in much reporting despite their distinct legal definitions, with consequences for how the legitimacy of claims was publicly judged. Crawley and Skleparis (2018) have since shown how the same categorical slippage shaped the politics of the 2015–2016 "refugee crisis" in Europe, and such representations intensify further around moments of political contest (KhosraviNik, 2010; Krzyżanowski *et al.*, 2018).

The securitization–humanitarianism dichotomy has been useful for analytical purposes and is supported by empirical observations. It has enabled scholars to identify the links between editorial position and recurring words, to make systematic comparisons between outlets, and to identify the associations between media framing and public opinion and policy orientations (Eberl *et al.*, 2018). Its explanatory reach, however, is limited, because it treats outlet ideology as the main source of variation and pays comparatively little attention to when coverage appears and whom it concerns. The dichotomy cannot readily explain why outlets whose baseline profile is securitizing extended an unqualified welcome to Ukrainian refugees in 2022, or why outlets whose baseline profile is humanitarian adopted crisis vocabulary when reporting Channel crossings. Recent research has begun to address this problem directly. Georgiou and Zaborowski (2017) and Chouliaraki and Georgiou (2022) show that hierarchies of deservingness track perceived cultural and racial proximity and that these distinctions hold across editorial alignments. Holmes and Castañeda (2016) identify the same sorting in German coverage, Krzyżanowski *et al.* (2018) situate it within the wider politicization of the 2015 "refugee crisis," and Wigger *et al.* (2022) document liberal German outlets adopting racialized and criminalizing repertoires after the Cologne assaults. Studies of the 2022 displacement from Ukraine point in the same direction: Ukrainian refugees were constructed in early news coverage as vulnerable and only contingently displaced (Sambaraju & Shrikant, 2023), and public and official responses to them differed markedly from those extended to Syrian and Afghan refugees (De Coninck, 2023; Garland & Lee, 2025). Research published in this journal likewise shows that the reception of displaced people depends on how particular populations are perceived and categorized, whether in public opinion toward Rohingya refugees in India (Verma & Kironska, 2024), in the political recognition of refugees in Brazilian cities (de Oliveira, 2023), or in the pressures on cultural identity faced by forced migrants more generally (Bhat *et al.*, 2025). This collection of work points to the possibility that the identity of migrants may impact as much on framing as the ideology of the outlet. However, that may not be testable with the single-event designs in which the identity of migrants must remain fixed. These

deviations therefore require attention to factors beyond outlet ideology—particularly which populations are being covered and when the coverage occurs.

1.2. Corpus linguistics, critical discourse analysis, and existing evidence

The social is constitutive for critical discourse analysis (CDA), meaning that texts are not just representations of an already existing social reality, but are also involved in the production of that reality, and they can naturalize it. The concept of an order of discourse (Fairclough, 2015) is the concept that media texts can be seen as being part of a network of institutional practice, and that any particular article both embodies and reinforces a conventional discourse on the article's topic. The discourse-historical approach adds a diachronic dimension by requiring that discursive strategies be interpreted within the historical and political context in which they occur (Reisigl & Wodak, 2016; Wodak & Meyer, 2016). This requirement is directly relevant to a study spanning a decade of policy and political change, and the approach's strategies of nomination, predication, and mitigation or intensification inform the coding categories described in Section 2.3.

Thus, ideology works most effectively under cover, i.e., where it is least seen and thus least susceptible to the sort of professional checks that are normally done by a news organization. This is not to say that ideology refers to a collection of political positions adopted by an organization; rather, it is in the sense of Fairclough (2015) and van Dijk (2018), as a set of taken-for-granted assumptions that are naturalized through repetition in many linguistic instances and that sustain certain relations of power among social groups. In the context of migration, it holds that patterns of collocation and evaluation are not necessarily an indicator of the ideological beliefs of individual journalists and/or of a coherent institutional voice, but rather are the product of ideological work that gets done in texts. The distinction matters. There is nothing inherently wrong with the combination of illegal and migrant, only the ubiquity of it, which makes it normal. Since then, these foundations have been carried over to other modes and platforms. Machin and Mayr (2012) speak about the ideological dimension of visual and typographic resources, and KhosraviNik (2017) examines the strengthening of the exclusionist migration discourse in digital media and the evolution of the relationship between elite and vernacular discourse. The framework is also transferable between contexts. It is used by Yu and Brookes (2024) in the context of Chinese and British coverage of coronavirus disease 2019 (COVID-19) and by Hu and Cai (2025) in the coverage of left-behind children in English-language publications, demonstrating how seemingly "positive" language can be used for ideological ends such as presenting constraint and limiting individual agency. This final discovery illuminates migration studies, as it warns against taking humanitarian framing as a given emancipatory approach.

One methodological criticism of CDA is that analysts may selectively choose examples that support their preconceived interpretations. Corpus linguistics (CL) addresses this objection directly by identifying what is typical across collections of text too large for close reading, and by locating patterns that are salient in the data rather than to the reader (Baker, 2010; McEnery & Hardie, 2012). Its central construct is collocation. Similar words acquire attitudinal coloring from the company they keep; the evaluative residue of repeated co-occurrence is largely unnoticed by language users, but can be readily measured on a large scale (Hunston, 2022). Lexical priming theory is a theory that helps account for why such associations are important for readers: the repeated exposure of a word in a context creates expectations that will be automatically triggered on subsequent mentions of the word (Hoey, 2005). Baker *et al.* (2008) formulated, in the case of their examination of the treatment of refugees and asylum seekers in the UK press, the connection between corpus and critical approaches. The synergy is expressed on both fronts: corpus techniques provide representativeness and minimize the risk of cherry-picking, while critical analysis offers the interpretive framework, which explains the social significance of the pattern. Partington *et al.* (2013) encapsulated the method in corpus-assisted discourse studies, and Baker (2023) produced a recent statement of the methods. Since then, Bednarek *et al.* (2024) have made an appeal for more transparency in the procedure and for developing an accountability framework on the basis of replicability, reflexive analysis, and cross-validation, which already forms the basis for reporting, as described in Section 2.

The approach has been applied widely. Liu (2024) shows how ideological nuances are conveyed in coverage of the 2023 Israel–Hamas war; Nairi (2025) examines how Tunisian media link sub-Saharan African migrants to crime and economic hardship; Islentyeva (2020) provides a systematic study of ideological patterning in the British press; and Del Fante (2018, 2024) compares UK and Italian coverage and proposes a broader agenda for corpus-assisted critical discourse studies of migration. Nartey and Mwinlaaru's (2019) meta-analysis of a decade of CL + CDA research confirms the effectiveness of the combined approach while identifying design constraints that are taken up in Section 1.3.

The comparative evidence indicates that ideology is mediated, and not an alternative to, national journalistic

conventions. The UK and the US share the Liberal Model of media systems, but there is one significant difference: the UK press has remained much more openly partisan, while the US has developed a professional norm of objectivity (Hallin & Mancini, 2004, 2017). It is the same with the role orientations: Hanusch and Hanitzsch (2017) identify a higher degree of leaning toward advocacy roles for UK journalists, and detachment roles for US journalists. There are reasons to believe, however, that ideological positioning in US coverage would work predominantly on the lexicological and source-selection level, instead of through overt evaluation. According to discursive institutionalism, ideas and discourse can be transmitted by institutions and replicated through institutional practices, which can also be unintentional, and that is how the longevity of such conventions is explained (Schmidt, 2008). Thus, the present analysis can examine whether systematic framing can occur under such conditions or not, as a declaration of "objectivity" could cause framing to become more difficult to observe, but not necessarily less or unreal.

1.3. Research gap, study objectives, and research questions

The literature available presents ideas with three limitations. First, the number of systematic inter-country comparisons of Anglophone media systems, based on an integrated CL + CDA approach, is still limited. Nartey and Mwinlaaru (2019) indicate that designs are mostly national and intra-European, and that cross-national work is challenging due to the difficulty of comparing corpora across national contexts. An established UK baseline is provided by Gabrielatos and Baker (2008), although their study did not include US media coverage. Second, longitudinal designs are scarce. Most studies focus on a single event or a short span of time, so it remains unclear whether the patterns they uncover reflect peculiarities of the period sampled or enduring features of outlets (Brezina, 2018; Kwak *et al.*, 2020; Taylor, 2021). Third, and as a consequence of the first two, the relationship between event-driven and ideological variation has not been examined within a single dataset. Evidence of stable ideological divergence and evidence of convergence around particular crises have accumulated largely in parallel and have been treated as competing explanations of the same process. A cross-national, cross-ideological, and longitudinal design is needed to establish how the balance between them shifts.

The UK and the US form an analytically useful pair. Both are Anglophone liberal democracies whose commercially financed and politically differentiated national news systems belong to the same model in Hallin and Mancini's (2004) typology, and both experienced a rise in populist anti-immigration mobilization between 2015 and 2024. Their migration geographies, however, differ considerably. The UK debate has centered on free movement within Europe, post-Brexit policy reconstruction, and Channel crossings (Consterdine & Hampshire, 2014; Mayblin *et al.*, 2024), whereas the US debate has centered on migration from Latin America, border militarization, and the status of unauthorized residents (Chavez, 2021). A comparison of two similar media systems—but subject to different migration contexts—can isolate the patterns of ideological framing to a greater extent than can be done by comparing different regulatory traditions. No translation is necessary as it is the same language for the two systems, and lexical/collocational patterns can be compared directly.

The study is significant in three respects. Empirically, it examines whether patterns documented in the UK press extend to a structurally similar system with a different migration geography, and whether they hold across a decade of substantial political change. Conceptually, it tests ideology-centered explanations of migration framing against evidence that framing varies with events and with the identity of the populations concerned. Practically, it informs media literacy education and debates about migration reporting, because its results indicate that evaluative framing operates less through explicit editorializing than through collocation and grammatical voice, features that conventional bias detection does not capture. The design also has a cost. The analysis is limited to two high-income Anglophone systems and has little to say about migration framing in sending-country media or in the Global South, where most of the world's displaced people live and where different representational logics prevail (Ullah & Chattoraj, 2023; Verma & Kironska, 2024).

This study examines the ideological framing of migration in UK and US national news outlets from 2015 to 2024. Combining CL and CDA, it identifies the lexical, collocational, and syntactic patterns surrounding the lemmas "migrant," "refugee," and "immigrant," and it interprets the discursive strategies (lexicalization, mitigation, presupposition, metaphor, and agency attribution) through which migrants are constructed as threats, as victims, or as rights-bearing subjects. All three lemmas are profiled with the same frequency, collocation, polarity, and agency procedures (Section 2.3). The study addresses one principal research question (RQ) and two sub-RQs:

(i) RQ1: How is migration ideologically framed in UK and US national news outlets between 2015 and 2024?
(ii) RQ1a: How do left-leaning and right-leaning outlets differ in their lexical, metaphorical, and evaluative framing of migrants, refugees, and immigrants?
(iii) RQ1b: How do these framings evolve over time in

relation to major socio-political events, including the 2015 European refugee crisis, the Brexit referendum, changes in US immigration policy, and the 2022 Ukrainian refugee crisis?

2. Data and methods

2.1. Corpus compilation and construction

The study draws on a purpose-built corpus of 500 news articles published between January 1, 2015, and December 31, 2024, by four English-language outlets: *The Guardian* and *The Telegraph* in the UK, and *The New York Times* and Fox News in the US. Outlets were selected to represent contrasting ideological positions within each national system, following prior classifications (Islentyeva, 2020; Jurkowitz *et al.*, 2020; Phelan, 2022): *The Guardian* and *The New York Times* as left-leaning, and *The Telegraph* and Fox News as right-leaning. The 10-year timeframe was selected since it had seen a particularly turbulent discourse, such as the European refugee crisis in 2015, Brexit in 2016 and its implementation, several changes in US immigration policies, increased rhetoric of populist anti-immigration, the COVID-19 pandemic, the withdrawal from Afghanistan in 2021 and the Ukrainian refugee crisis in 2022.

Such labels are not assumed to be fixed characteristics of the organizations in question, but are analytic tools used to get a handle on ideology. They suggest that they are evidence of an outlet's overall output during a specific period of time, not that every article or journalist from that outlet has this uniform stance. The evidence in Section 3 indicates that there is considerable variation within outlets, which supports this qualified usage.

Fox News is not the same institution as the other three, as it is a cable television channel instead of a newspaper. The Fox News material includes only written articles that appeared on foxnews.com—broadcast transcripts and wire-service copy (e.g., Associated Press) that appeared on the site were not included. Fox News was included because it is the most popular source of news among US conservatives and a key player in the right-wing media universe (Benkler *et al.*, 2018; Jurkowitz *et al.*, 2020). Because digital articles produced by a broadcaster may be more reliant on quotation than newspaper articles, all frequencies are normalized, the same genre quotas are applied to all outlets, and quoted speech is distinguished from editorial voice (Section 2.4).

Articles were retrieved from the outlets' own digital archives and from ProQuest News & Newspapers using search terms targeting the lemmas "migrant*," "refugee*," "immigrant*," "asylum seeker*," and "border*." An article was included if a search term appeared in the headline or lead paragraph and if migration policy, an asylum or refugee event, or border control was its main focus. Articles were excluded if the search term was used only incidentally (for example, in an unrelated news report, a letters page, or a listing) or if the full text could not be obtained. Duplicate and syndicated articles were removed.

2.1.1. Temporal allocation

Articles were allocated across the decade by a design that combined an even baseline with additional sampling around major events. The baseline comprised ten articles per outlet per year (400 articles); a further 100 articles were allocated to years, and in some cases to one country's outlets only, in which major migration-related events occurred. For example, each UK outlet received eight additional articles in 2016, the year of the Brexit referendum, compared with two for each US outlet, and each US outlet received five additional articles in 2018, the year of the family separation policy. The purpose of event weighting was to secure sufficient material from the periods in which migration discourse was most intensively contested. Its consequence is that the corpus does not reproduce the actual volume of coverage from year to year, so annual article counts are not interpreted as indicators of coverage intensity. Table 1 reports the allocation by outlet, year, and period.

2.1.2. Genre stratification

Within each outlet, articles were stratified by year and genre (news report, editorial/opinion, and feature) and randomly selected from the eligible articles within each year–genre stratum. Genre quotas were applied separately to each outlet across the entire study period to ensure that news reports constituted the largest portion of each subcorpus while retaining sufficient explicitly evaluative material for stance analysis. Genre classifications were based on publication metadata, such as section labels (e.g., "Opinion"), and were manually verified. Table 2 presents the distribution of articles by outlet and genre, together with the mean article length and type–token ratio for each subcorpus. The final corpus comprises approximately 1.2 million words.

Table 1. Temporal distribution of the corpus by outlet, year, and analytical period

Period	Year	GU	TE	NYT	FN	Total	Key events
P1	2015	15	15	15	15	60	European refugee crisis
P1	2016	18	18	12	12	60	Brexit referendum; US presidential campaign
P2	2017	12	12	13	13	50	US executive orders on immigration
P2	2018	10	10	15	15	50	US family separation policy
P2	2019	10	10	10	10	40	UK general election
P3	2020	13	13	12	12	50	COVID-19 pandemic
P3	2021	12	12	13	13	50	Withdrawal from Afghanistan
P4	2022	12	12	13	13	50	Ukrainian refugee crisis
P4	2023	12	12	11	11	46	Small-boat crossings (UK)
P4	2024	11	11	11	11	44	US presidential election
Total		125	125	125	125	500	

Notes: GU = *The Guardian*; TE = *The Telegraph*; NYT = *The New York Times*; FN = Fox News. Articles per period (P1/P2/P3/P4): GU and TE, 33/32/25/35; NYT and FN, 27/38/25/35; all outlets, 120/140/100/140.
Abbreviations: UK: United Kingdom; US: United States.

Table 2. Corpus composition by outlet and genre, with mean article length and type–token ratio

Subcorpus	News reports	Editorial/opinion	Feature	Total articles	Mean article length (words)	Type–token ratio
The Guardian	72	24	29	125	2,240	0.31
The Telegraph	81	26	18	125	2,560	0.33
The New York Times	75	29	21	125	2,320	0.27
Fox News	82	21	22	125	2,480	0.25
Total	310	100	90	500	2,400	0.29

2.1.3. Sample size and the scope of diachronic claims

On average, the corpus contains 12.5 articles per outlet per year, which is a thin slice of each outlet's annual output. The corpus is therefore not designed to estimate how much any outlet published about migration, and it does not support year-by-year or genre-by-year comparisons: a single year–genre cell contains, on average, only three to five articles per outlet. Three decisions follow from this constraint, and they govern every quantitative claim in Section 3. First, diachronic comparison is conducted across four periods rather than single years: P1 (2015–2016), P2 (2017–2019), P3 (2020–2021), and P4 (2022–2024). Each outlet is represented by 25–38 sampled articles per period (about 60,000 to 90,000 words). Second, the unit of quantitative evidence is not the article but the occurrence of a node lemma in its co-text. Each comparison rests on the number of occurrences of each target lemma, disaggregated by outlet and by period (Section 3.1), and a collocate is retained only if it occurs in at least three different articles, so that no collocation rests on a single article. Third, genre is not treated as a comparative variable: stratification ensures that each subcorpus contains comparable proportions of news reports, opinion pieces, and features, and no claims are made about genre differences within particular years.

While small with regard to the conventions of modern CL, the corpus size is similar to other corpus-assisted discourse analyses that combine quantitative methods with close qualitative reading (Partington *et al.*, 2013). Its suitability was evaluated in light of the goal of the analysis, and not in terms of a general size criterion: The three node

lemmas are used in all of the outlet subcorpora, and in each period, and thus their collocational profiles can be compared; low-frequency phenomena are described in a descriptive way and used with caution.

To prepare the articles for analysis, boilerplate (navigation text, image captions, related-article links, and advertising labels) and non-alphabetic "noise" were removed from the articles, but repeated whitespace was collapsed, quotation marks were standardized, and the articles were not otherwise edited.

The four national and mainstream outlets were deliberately selected to enable controlled ideological and cross-national comparison. This does not include local, regional, or tabloid newspapers, broadcast news programming, or digital-native publications, which can be different in nature in terms of how they portray migration and can have significant readerships. The implications of this limitation in terms of generalizability are discussed in Section 4.7.

2.1.4. Terminology

For this article, the authors employ the term "undocumented" (and, when necessary, "irregular") to describe people who lack legal authorization to enter or remain in a country, in accordance with the terminology recommended by the International Organization for Migration (2019). The term "illegal" is used only when referring to language employed by the news outlets and is enclosed in double quotation marks when presented as a corpus item or quoted directly from the outlet. The distinction between using a term to describe people and mentioning it as an object of analysis is central to CDA. The decision to avoid "illegal" as a descriptive label also reflects the study's finding that the term attributes a property of administrative status to the person (Section 3.5).

2.2. Analytical tools

Two corpus analysis packages were used in combination. AntConc (Anthony, 2022) was used for frequency profiling and keyword-in-context (KWIC) concordance analysis, generating frequency lists for migration-related terms and enabling qualitative reading of the immediate co-text of target items. Collocation statistics, dispersion measurement, and visualization were performed using LancsBox (Brezina *et al.*, 2021), including the GraphColl collocation networks (Gablasova *et al.*, 2017), dispersion plots, and subcorpus comparison between the ideological and national division of the data set (Baker *et al.*, 2008).

The two tools enabled cross-validation of results and were not redundant. Measures of frequency and collocation produced in one package were verified against the other, with a comparison of discrepancies arising primarily as a result of differences in default lemmatization and tokenization parameters; before these were resolved, the analysis proceeded by standardizing parameters across platforms. The use of the combination also enabled close qualitative reading of the concordance lines to go together with quantitative measures of association, which is not easily possible with either tool alone (Hart, 2020; Wodak & Meyer, 2016). The data were coded using qualitative coding in NVivo (Version 14, Lumivero, United States of America).

2.3. Stages of analysis

The analysis proceeded in three stages, and each stage was applied in the same way to all three node lemmas: "migrant," "refugee," and "immigrant" (each including its plural form).

(i) Stage 1: Frequency profiling. Raw and normalized frequencies (per 100,000 words) of the three lemmas were calculated for the full corpus, for each outlet, for each ideological group, and for each period. Differences between subcorpora were tested with log-likelihood (LL), and their size was expressed as log ratio (Gabrielatos, 2018). This stage establishes differences in lexical preference, notably the preference for "migrant" over "refugee" in right-leaning outlets, which withholds the legal and moral status that the refugee designation confers (Keyel, 2024; KhosraviNik, 2010), as well as shifts in term preference across the four periods. The results are reported in Section 3.1.

(ii) Stage 2: Collocation. Collocates of each lemma were extracted separately within a 5L–5R window (i.e., five words to the left and five words to the right of the node word). A candidate collocate had to meet four criteria jointly: (a) at least five co-occurrences with the node; (b) mutual information (MI) ≥ 3; (c) LL ≥ 15.13, corresponding to $p < 0.0001$ (Rayson *et al.*, 2004); and (d) occurrence in at least three different texts. MI favors exclusive but infrequent collocates and LL favors frequent ones, so requiring both reduces the bias of either measure (Evert, 2008; Gablasova *et al.*, 2017). Semantic fields around each node were mapped with GraphColl networks, and the same procedure was applied within each period to trace changes in those fields.

Attributive forms (e.g., illegal, undocumented) were not considered node words. They were extracted as collocates of the node lemmas immigrant and migrant, in the L1 (immediately premodifying) position, thus forming the noun phrases "illegal immigrant(s)," "illegal migrant(s)," "undocumented immigrant(s)," and "undocumented

migrant(s)." The concordance lines were then read manually to ensure that only attributive uses modifying the node had been counted; those uses that do not modify a node lemma, such as illegal crossing or undocumented status, were excluded from the counts reported in Section 3.5.1, but not the co-text. This procedure ties the analysis of the illegal–undocumented alternation directly to the immigrant lemma, where the alternation is concentrated, while recording the parallel forms with migrant for comparison.

Metaphors were identified by reading all concordance lines of the three node lemmas and of *migration* and *immigration* for metaphorical expressions in their co-text (Stefanowitsch, 2006). Each candidate was tested with Metaphor Identification Procedure Vrije Universiteit (MIPVU) (Steen *et al.*, 2010): an expression was coded as metaphor-related when its contextual meaning contrasted with a more basic meaning, and metaphor-related expressions were then grouped by conceptual source domain (Lakoff & Johnson, 2003).

(iii) Stage 3: Qualitative coding. Qualitative thematic coding was applied in NVivo to KWIC lines of all three lemmas and their surrounding context, following established CDA frameworks (Fairclough, 2015; Reisigl & Wodak, 2016). Each category was operationalized with an explicit definition and at least two corpus-derived examples recorded in a coding manual developed during a pilot round on 5% of the corpus. Lexicalization captured evaluatively loaded referential choices (e.g., "hordes," "victims," and "job-seekers"). Mitigation and intensification captured hedging and boosting in descriptions of migration, including hedges such as "some migrants" and boosters such as "flooded" and "overwhelmingly." Presupposition captured implicit assumptions embedded in constructions, particularly those positioning migration as inherently burdensome or threatening.

Evaluative polarity was coded at the level of the individual evaluative item, defined as an adjective, noun, verb, or adverb within the 5L–5R span of a node lemma that attributes a quality to the people referred to, to their movement, or to their presence in the receiving society. Candidate items were drawn from the Stage 2 collocate lists and from KWIC reading. The classification was grounded in the ATTITUDE system of appraisal theory, chiefly judgment (of people and their conduct) and appreciation (of phenomena), together with the concept of semantic prosody (Hunston, 2022; Martin & White, 2005). Table 3 presents the criteria. Polarity was assigned by contextual meaning rather than by dictionary lookup: each item was read in its full sentence, and items whose evaluation was reversed or disowned in context (for example, migrants are not a threat, or a label explicitly rejected by the writer) were coded by their contextual orientation. Items that were evaluatively neutral (e.g., Syrian, policy, border) or whose orientation could not be determined were excluded. Items of sympathetic victimhood (e.g., vulnerable, desperate) were coded as positive, because they evaluate the person's claim to protection favorably; this coding records the direction of evaluation only, and Section 4.2 discusses why such framing is not necessarily emancipatory. To compare outlets on an equal footing, 50 coded items per outlet and per lemma were drawn by random sampling from the pool of coded items, balanced across the four periods. Quotation status was recorded for every item.

Two aspects of agency were coded. For each reference to an enforcement action (removal, detention, interdiction, or separation), coders recorded whether the responsible institutional actor was expressed, backgrounded, or omitted. For each node lemma, coders also recorded agency-attributing verbs, defined as verbs of volitional action (e.g., "flee," "seek," "apply," "work," and "build") of which the node is the grammatical subject and the semantic actor.

Table 3. Coding criteria for evaluative polarity

Polarity	Operational criterion	Typical items in the corpus
Positive	The item evaluates the people referred to as legitimate, capable, contributing, or entitled to protection, or evaluates their movement or presence as beneficial or deserving of support	"resilient," "contributor," "humanitarian," "welcome," "essential," "vulnerable"
Negative	The item evaluates the people referred to as a threat, burden, criminal, or illegitimate presence, or evaluates their movement as an uncontrolled mass or danger	"illegal," "burden," "uncontrolled," "invasion," "threat," "criminal"
Excluded	The item is evaluatively neutral, or its orientation cannot be determined from the sentence and the surrounding paragraph	"Syrian," "policy," "border," "processing"

2.4. Validity and reliability

Several procedures were adopted to ensure the reliability of the analysis. Intercoder agreement was assessed by having a second researcher independently code 10% of the corpus, which yielded a Cohen's kappa of 0.82, indicating strong agreement. Agreement for evaluative polarity, assessed on the same subsample, was $\kappa = 0.80$. Disagreements were first resolved through discussion with reference to the coding manual; any that remained were adjudicated by a third member of the research team, and the coding manual was revised to clarify the relevant category boundary before coding continued.

Quantitative results are reported with their corresponding counts and association indices and, where applicable, are tested statistically rather than judged visually (Sections 3.3 and 3.6). MI and LL values were computed under the standardized parameters described in Section 2.2, and collocational claims rest on those values together with the thresholds set out in Section 2.3. Where a feature occurs fewer than 10 times in a subcorpus, results are reported descriptively and are not used to support comparative claims. Concordance examples reproduced in Section 3 are verbatim; ellipsis points mark abridgment, and outlet and date are given in parentheses after each example.

A persistent validity question in media corpus analysis concerns language quoted from sources rather than written in the outlet's own voice. All concordance lines examined qualitatively were therefore coded for quotation status (direct quotation, indirect reported speech, or editorial voice), and the results are reported in Section 3.5. The transparency measures set out in this section specify what is required for replication, in line with the accountability standards proposed by Bednarek *et al.* (2024).

3. Results

Sections 3.1 to 3.3 report the frequency, collocation, and evaluative profiles of the three node lemmas. Section 3.4 traces change across the four periods, and Sections 3.5 to 3.7 report the concordance, agency, and metaphor analyses.

3.1. Frequency profiles of the three node lemmas

Table 4 reports the raw and normalized frequencies of "migrant," "refugee," and "immigrant" in each outlet subcorpus. Across the full decade, the left-leaning outlets used refugee more often, whereas the right-leaning outlets preferred migrant; both contrasts were substantial (refugee: LL = 83.01, log ratio = +0.67; migrant: LL = 131.05, log ratio = −0.75). Immigrant was concentrated in the US subcorpus but showed almost no aggregate left–right difference (LL = 0.12, log ratio = +0.02), indicating that national migration geography mattered more than ideology for this label.

Table 5 reports the same frequencies by ideological group and period. It shows the evidential base for the diachronic analysis in Section 3.4.

Table 4. Raw frequencies (normalized per 100,000 words in parentheses) of the node lemmas, by outlet

Outlet	"migrant(s)"	"refugee(s)"	"immigrant(s)"
The Guardian	310 (110.7)	488 (174.3)	252 (90.0)
The Telegraph	602 (188.1)	295 (92.2)	221 (69.1)
The New York Times	400 (137.9)	423 (145.9)	524 (180.7)
Fox News	720 (232.3)	336 (108.4)	622 (200.6)
Total	2,032 (169.3)	1,542 (128.5)	1,619 (134.9)

Notes: Left- vs right-leaning comparison: migrant log-likelihood (LL) = 131.05, log ratio = −0.75; refugee LL = 83.01, log ratio = +0.67; immigrant LL = 0.12, log ratio = +0.02.

Table 5. Normalized frequencies (per 100,000 words; raw counts in parentheses) of the node lemmas, by ideological group and period

Period	"migrant" L	"migrant" R	"refugee" L	"refugee" R	"immigrant" L	"immigrant" R
P1 (2015–2016)	138 (174)	206 (272)	211 (266)	153 (202)	113 (137)	127 (167)
P2 (2017–2019)	119 (150)	207 (273)	117 (148)	76 (100)	151 (183)	179 (233)
P3 (2020–2021)	91 (91)	121 (121)	144 (144)	120 (120)	118 (114)	132 (129)
P4 (2022–2024)	123 (295)	273 (656)	147 (353)	87 (209)	139 (334)	137 (322)

Note: L = Left-leaning outlets; R = Right-leaning outlets.

3.2. Collocational patterns

Collocation analysis produced markedly different semantic profiles across the four outlets for each of the three lemmas. Table 6 reports the five strongest collocates of each lemma in each outlet, identified with the criteria set out in Section 2.3.

3.2.1. Migrant

The four profiles differ less in their degree of negativity than in the object of concern around which each organizes migration. *The Guardian*'s collocates construct migration as something that migrants do and are entitled to: "journey" denotes movement undertaken by the migrant, while "rights" and "asylum" denote entitlements under international law. Only "crisis," a collocation that is also frequent in *The Telegraph* (Section 3.5), names the phenomenon as an emergency, and in *The Guardian*'s usage it attaches to the systems that fail migrants rather than to their arrival. *The Telegraph*'s profile organizes migration around state capacity. "Influx" and "threat" construct arrival as a quantity exceeding tolerance, while "control" and "security" specify administrative containment as the appropriate response. In contrast to Fox News, this is analytically important, and not much has been studied in previous research. Fox News clusters around "crime"/"criminality" and "invasion"/"confrontation" ("wall"/"deportation"), while *The Telegraph* ones cluster around "governance"/"capacity." Both outlets securitize, but *The Telegraph* in a language of managerial tension instead of criminal menace. This register suits a broadsheet model in which restrictionist positions are presented as sensible policy, and readers are less likely to recognize its evaluative elements.

A parallel distinction separates the two left-leaning outlets. *The Guardian*'s profile centers on the migrant as a rights-bearer and as a person with experiences, whereas *The New York Times*' profile centers on migration as an object of policy. "Policy," "reform," and "undocumented" place the phenomenon within legislative and administrative processes rather than at the margins of a receiving society. This is the collocational signature that Hallin and Mancini's (2004) account of professional "objectivity" in US journalism would lead one to expect: a sympathetic stance conveyed through procedural and structural vocabulary rather than through affective terms. The clearest example is the choice of "undocumented" over "illegal." "Illegal" characterizes the person, whereas "undocumented" characterizes their administrative status (De Genova, 2013).

3.2.2. Refugee

The refugee profiles are more consistently humanitarian, in contrast to migrant profiles. *The Telegraph* and Fox News are most likely to associate refugees with Ukrainians and Afghans singled out as particularly deserving of protection; *The Guardian* and *The New York Times* more often link refugees to asylum, protection, resettlement, family, humanitarian need. The securitizing of vocabulary is restricted in all four refugee profiles and takes place instead with the use of the node migrant, which suggests that the choice of node does part of the work of evaluation before any explicit predication is formulated (Gabrielatos & Baker, 2008; Keyel, 2024).

3.2.3. Immigrant

Immigrant is focused on the US subcorpus and is mostly part of the discussions about illegal stay, legalization, work, borders, etc. (Chavez, 2021). While the lexical opposition between illegal and undocumented is apparent in the L1 counts reported in Section 3.5.1, *The New York Times* profile is arranged around the topic of undocumented status, workers, community, reform, and contribution, while the Fox News profile is set against the background

Table 6. Five strongest collocates of the node lemmas, by outlet

Outlet	"migrant"	"refugee"	"immigrant"
The Guardian	crisis, journey, rights, asylum, border	asylum, protection, resettlement, family, humanitarian	integration, community, work, citizenship, settled
The Telegraph	influx, threat, illegal, control, security	Ukrainian, Afghan, protection, family, resettlement	integration, work, citizenship, settled, community
The New York Times	policy, reform, border, undocumented, community	resettlement, asylum, family, protection, humanitarian	undocumented, workers, community, reform, contribution
Fox News	invasion, crime, illegal, wall, deportation	Ukrainian, Afghan, family, welcome, protection	illegal, border, crime, deportation, enforcement

Note: 5L–5R; frequency ≥ 5, mutual information ≥ 3, log-likelihood ≥ 15.13, ≥3 texts.

of illegal aspects of status, workers, community, reform, and contribution. In the UK outlets, immigrant occurs less often and is used mainly for settled populations, integration, work, and citizenship rather than for newly arriving groups, which are more often named migrants or refugees.

These are aggregate profiles with an important limitation. The patterns in Table 6 are those of the full 10-year corpus, and the results in Section 3.4 indicate that they did not remain constant. Some events, most prominently the arrival of Ukrainian refugees in 2022, produced a convergence across outlets that is not visible in the aggregate, cross-sectional profiles.

3.3. Evaluative polarity

Evaluative polarity was coded with the criteria in Table 3 (Section 2.3). Figure 1 compares the distribution of positive and negative evaluative items collocating with "migrant" across the four outlets, and Table 7 reports the underlying counts for all three lemmas.

For migrants, the association between outlet and evaluative polarity is statistically significant and strong (Cramér's $V = 0.52$). The standardized residuals identify the two largest contributors to the overall statistic: *The Guardian*, with a residual above +3, and Fox News, with a residual below −3. The intermediate values for *The Telegraph* and *The New York Times* are also informative. Both depart from the corpus mean in the same direction as the outlet that shares their ideological position (Fox News and *The Guardian*, respectively), but by a smaller margin. This suggests that the explicitness of evaluation depends on register as well as on ideological position: these two outlets moderate the expression of evaluation without changing its direction. Outlet differences are smaller for "refugee" than for "migrant" ($V = 0.26$), because even the right-leaning outlets evaluate refugees positively when the populations concerned are represented as deserving. The difference widens again for "immigrant" ($V = 0.40$), driven chiefly by the contrast between the positive economic and community vocabulary of *The New York Times* and the negative illegality and enforcement vocabulary of Fox News.

3.4. Temporal evolution of migration discourse, 2015–2024

Because each outlet contributes 25–38 articles to each period (Table 1), the diachronic analysis compares periods rather than years, and the dated concordance lines cited below illustrate period-level patterns rather than establishing year-level trends (Section 2.1). Read in this way, the analysis indicates that lexical preference and evaluative orientation both changed across the decade, in ways associated with major socio-political events.

(i) P1 (2015–2016). During the European refugee crisis, the Brexit referendum, and the US presidential campaign, *refugee* was frequent in all four outlets (Table 5). Even so, the outlets differed in how they named the people concerned: *The Guardian* and *The New York Times* used "refugee," whereas *The Telegraph* and Fox News preferred "migrant" and "illegal immigrant." The difference is not merely stylistic. "Refugee" carries a protective connotation that "migrant" lacks (Gabrielatos & Baker, 2008; Keyel, 2024), so the choice between them is a choice about obligation. The framing was polarized from the outset. *The Guardian* stressed Europe's "responsibility" toward people escaping violence in Syria (Guardian, September 2015), whereas *The Telegraph* warned that the influx threatened to overstretch European border control (Telegraph, September 2015). The referendum appears to have strengthened the association between migration and sovereignty in *The Telegraph*, where the control of numbers became a recurrent formulation (Telegraph, June 2016).

(ii) P2 (2017–2019). In the US subcorpus, "refugee" became less frequent and was displaced by "immigrant" and by the vocabulary of the border and domestic enforcement. The UK subcorpus shows a comparable decline, as the implementation of Brexit dominated political attention. The 2018 family separation policy produced a sharp contrast in the US subcorpus between humanizing detail in *The New York Times* and legalistic framing in Fox News, which described the processing of people crossing without authorization as the application of existing law (New York Times, June 2018; Fox News, June 2018).

(iii) P3 (2020–2021). Migration coverage receded during the first phase of the COVID-19 pandemic and returned with the withdrawal from Afghanistan, when "refugee" was again frequent in all four outlets, including those with securitizing baselines. In this period, Fox News adopted humanitarian language for Afghans who had supported US forces, holding that "Afghan refugees who aided US forces deserve protection" (Fox News, August 2021). The formulation grounded entitlement not in need but in service rendered. In the same period, *The Guardian* described Channel crossings in a vocabulary of capacity and crisis: "Small boats crisis tests UK asylum system capacity" (Guardian, November 2021).

(iv) P4 (2022–2024). The most striking convergence was the humanizing representation of Ukrainian refugees in 2022 across all four outlets (Section 4.3).

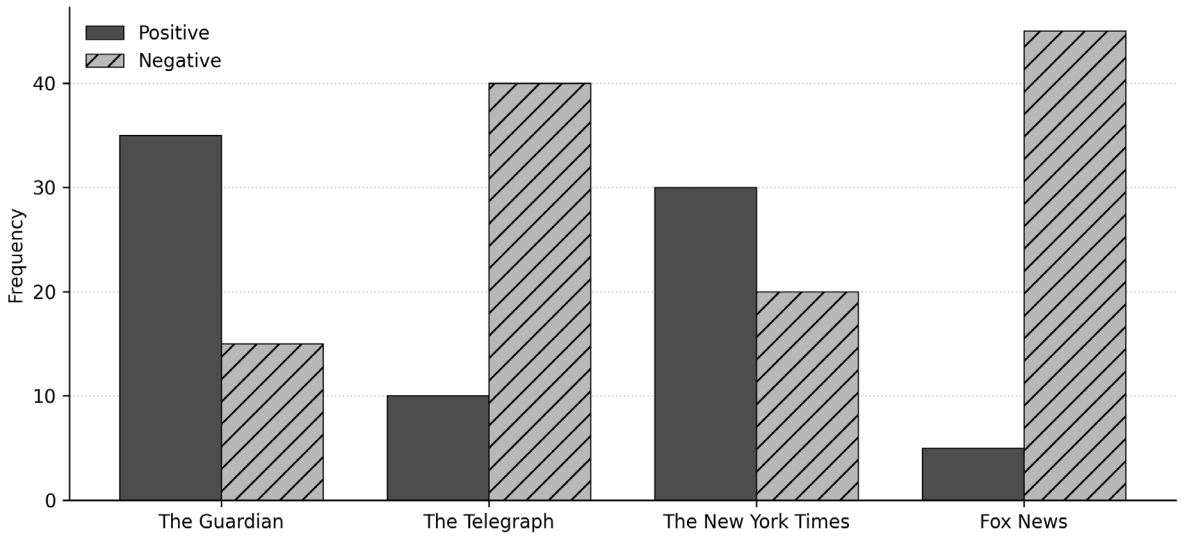


Figure 1. Frequency of positive and negative evaluative items collocating with "migrant," by outlet

Table 7. Distribution of evaluative polarity in items collocating with the node lemmas, by outlet

Lemma	Outlet	Positive	Negative	Total	% positive	Standardized residual (positive category)
"migrant"	*The Guardian*	35	15	50	70.0	+3.35
	The Telegraph	10	40	50	20.0	−2.24
	The New York Times	30	20	50	60.0	+2.24
	Fox News	5	45	50	10.0	−3.35
	Total	80	120	200	40.0	
"refugee"	*The Guardian*	39	11	50	78.0	+1.24
	The Telegraph	28	22	50	56.0	−0.71
	The New York Times	37	13	50	74.0	+0.88
	Fox News	24	26	50	48.0	−1.41
	Total	128	72	200	64.0	
"immigrant"	*The Guardian*	32	18	50	64.0	+1.75
	The Telegraph	18	32	50	36.0	−1.13
	The New York Times	34	16	50	68.0	+2.17
	Fox News	10	40	50	20.0	−2.78
	Total	94	106	200	47.0	

Note: Migrants: $\chi^2(3) = 54.17$, $p < 0.001$, Cramér's $V = 0.52$. Refugees: $\chi^2(3) = 13.37$, $p = 0.004$, $V = 0.26$. Immigrants: $\chi^2(3) = 31.71$, $p < 0.001$, $V = 0.40$.

By 2023–2024, securitizing vocabulary had returned, attached to the small-boats debate in the UK and to the election-year immigration debate in the US.

Collocational change across the periods points to a shift in the underlying construction of migration. In P1, migrant collocated predominantly with crisis, flood, and influx, and with refugee, asylum, and Syrian, forming an emergency register in which humanitarian and crisis vocabularies were intertwined. From P2 onward, the dominant collocates shifted toward illegal, border, and deportation, and, in the UK subcorpus, toward the cluster Channel, boat, and crossing. The movement is from a discourse of exceptional emergency toward one of permanent "illegality": migration ceases to be an event requiring a response and becomes a condition requiring enforcement.

The diachronic analysis also yields a comparative result. Baseline differences between left- and right-leaning outlets persisted throughout the decade. However, in the event-dominated periods (P3 and P4), the shifts within outlets between periods were of a magnitude comparable to the differences between outlets within a period. Given the size of the period subcorpora, this result is best read as evidence that event context can rival outlet ideology as a determinant of framing at particular moments, rather than as a precise estimate of their relative weight.

3.5. Concordance analysis

Frequency and collocation establish patterns at scale; concordance analysis shows how those patterns operate in actual discourse. This section reports a systematic concordance analysis of the principal collocations identified above.

The collocation "migrant crisis" occurred 147 times across the corpus, distributed unevenly and constructed differently by outlet. *The Guardian* (21 instances) blames the crisis on the failure of institutions. Phrases such as "The migrant crisis demands a coordinated European humanitarian response" and "Migrant crisis exposes failure of EU asylum system to protect vulnerable people" turn the problem away from that of the arrival, and "Behind the migrant crisis statistics are individuals fleeing persecution and war" explicitly counters the aggregation action the term performs. *The Telegraph* (38 instances) constructs the same collocation as external pressure on national capacity: "The migrant crisis threatens to undermine border security across the continent," "Migrant crisis: Government announces tougher enforcement measures," and "Public concern grows as migrant crisis continues unabated." That the collocation is nearly twice as frequent in *The Telegraph* is itself a finding, since it indicates differential salience as well as differential construction.

3.5.1. Immigrant: "Undocumented" versus "illegal"

As explained in Section 2.3, "illegal" and "undocumented" were analyzed as L1 premodifiers of the node lemmas "immigrant" and "migrant," not as independent node words. The alternation between them encodes position more directly than any other single variable in the corpus, and Table 8 shows that it is concentrated on "immigrant." Fox News used "illegal immigrant" or "illegal migrant" 89 times, whereas *The New York Times* used these forms 12 times and preferred "undocumented" (67 instances).

The grammatical environments differ as sharply as the terms. Fox News pairs "illegal" with law-enforcement verbs ("Illegal migrants apprehended at the border face immediate deportation proceedings," "Border patrol seizes illegal migrants attempting to evade detection") and with predications of harm: "The administration's failure to stop illegal migration endangers American communities." *The New York Times* pairs "undocumented" with economic and structural predicates: "Undocumented immigrants contribute billions annually to Social Security," "Advocates argue that undocumented status results from bureaucratic barriers," and "Undocumented workers essential to agricultural sectors face exploitation." The first construction locates the relevant property in the person and the appropriate response in enforcement; the second locates it in an administrative system and the appropriate response in reform. The contrast also explains why this article uses "undocumented" in its own voice (Section 2.1): the "illegal" construction does not describe a legal status so much as attribute deviance to a person.

3.5.2. Editorial voice and quotation

A recurring methodological concern in media discourse analysis is whether framing attributed to an outlet in fact originates with quoted political actors. Systematic coding of quotation status addressed this concern directly. Of the 23 instances of the "migrant" + "wall" collocation in Fox News, 14 occurred in direct quotation ("President Trump stated: 'We need a wall to stop the migrant invasion,' 'Republican senators argue the border wall will deter migrant flows'") and nine in editorial voice. The editorial instances retain securitizing vocabulary even where the political claim is reported neutrally or contested: "The proposed wall aims to control migrant entry points along the southern border" and "Critics of the wall say it won't stop determined migrants" both presuppose that migration is a flow requiring interception, the second doing so while ostensibly voicing opposition. The same pattern holds in reverse for *The Guardian*, where 19 of 31 instances of "migrant" + "rights" were attributed to advocacy sources such as Amnesty International and UN officials and

Table 8. The premodifiers illegal and undocumented in L1 position before the node lemmas immigrant and migrant, by outlet (raw frequencies)

Outlet	"illegal immigrant(s)"	"illegal migrant(s)"	"illegal" total	"undocumented immigrant(s)"	"undocumented migrant(s)"	"undocumented" total
The Guardian	4	6	10	29	8	37
The Telegraph	34	19	53	7	3	10
The New York Times	8	4	12	61	6	67
Fox News	72	17	89	9	2	11

12 occurred in editorial voice: "The policy threatens to undermine fundamental migrant rights protections" and "Migrant rights advocates warn of deteriorating conditions in detention facilities."

The distribution therefore does not suggest that outlets merely transmit the language of their preferred sources. In both cases, source selection and editorial voice work in the same direction, and the presuppositions carried by editorial formulations persist regardless of whose position is being reported.

3.6. Agency attribution

The voice in reporting of enforcement action is systematically different according to outlet, with grammatical voice varying. Right-leaning outlets frequently contain constructions where the institutional actor is not mentioned: "Migrants were removed from the country following visa violations" does not mention the immigration authorities that carried out the removal; "The detention facility was found to process asylum claims" does not mention the government that created it; "Border controls tightened in response to growing arrivals" does not mention the policymakers who tightened them, but rather frames this manipulation as a reaction. Left-leaning outlets are inclined to constructions that foreground and name the actor: "Home Office officials detained asylum seekers without legal representation," "ICE agents separated families at the border despite outcry from advocates," and "The administration implemented policies that systematically deny refugee protections."

The distribution of agency-attributing verbs (Section 2.3) is reported for "migrant" in Figure 2 and for all three lemmas in Table 9.

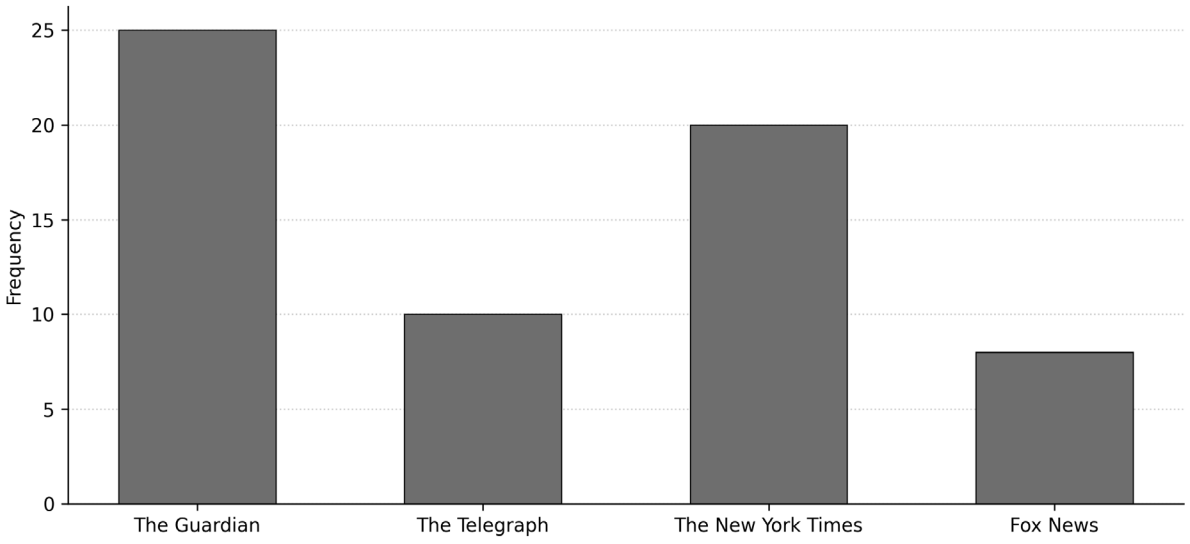


Figure 2. Distribution of agency-attributing verbs with "migrant(s)" as grammatical subject, by outlet

Table 9. Frequency of agency-attributing verbs with the node lemmas as grammatical subject, by outlet

Lemma	Outlet	Agency-attributing verbs	Rate per 100 articles	Standardized residual
	The Guardian	25	20.0	+2.33
	The Telegraph	10	8.0	−1.45
"migrant"	*The New York Times*	20	16.0	+1.07
	Fox News	8	6.4	−1.95
	Total	63	12.6	
	The Guardian	22	17.6	+1.21
	The Telegraph	14	11.2	−0.73
"refugee"	*The New York Times*	20	16.0	+0.73
	Fox News	12	9.6	−1.21
	Total	68	13.6	
	The Guardian	12	9.6	−0.72
	The Telegraph	9	7.2	−1.50
"immigrant"	*The New York Times*	28	22.4	+3.45
	Fox News	10	8.0	−1.24
	Total	59	11.8	

Note: Goodness-of-fit against a uniform distribution: migrant $\chi^2(3) = 12.49$, $p = 0.006$; refugee $\chi^2(3) = 4.00$, $p = 0.261$; immigrant $\chi^2(3) = 16.19$, $p = 0.001$.

For migrant, the distribution is far from even. *The Guardian* records an agency rate more than three times that of Fox News, with *The New York Times* and *The Telegraph* in intermediate positions, mirroring the ordering observed for evaluative polarity in Section 3.3. The pattern parallels that for enforcement: outlets that grammatically background the state actor also background the agency of migrants, constructing migration as a process in which no one in particular acts. Migrants undergo procedures that have no named authors. Left-leaning outlets, by contrast, attribute agency in both directions: to migrants who undertake journeys and to officials who detain them. Refugee is more evenly distributed and most often takes verbs of constrained movement, especially flee, escape, and seek, so agency remains bounded by displacement across all outlets. Immigrant shows a sharper outlet effect: economic agency verbs such as work, contribute, and build are concentrated in *The New York Times*, while the other outlets more often position immigrants as the objects of policy and enforcement.

3.7. Metaphorical framing

Metaphorical patterning follows similar lines and is similarly ideologically loaded. The metaphors most frequent in right-leaning outlets draw on the source domain of water: "A flood of migrants overwhelmed Mediterranean rescue operations," "The tide of immigration threatens to inundate border communities," and "Stemming the flow of illegal migration requires comprehensive border enforcement." This source domain constructs migration as a natural phenomenon without agency or intention, amenable only to physical containment (Charteris-Black, 2011; Taylor, 2021). Two entailments follow. Because floods have no agents, individual purpose is removed from the people who make up the flood. And because barriers are what one builds against floods, alternative responses are foreclosed at the level of conceptual structure rather than at the level of argument (Johnson *et al.*, 2024; Lakoff & Johnson, 2003).

Left-leaning outlets favor journey metaphors: "Migrants make dangerous journeys to reach safety and opportunity," "The journey to asylum is fraught with bureaucratic hurdles and challenges," and "Refugee families face difficult legal hurdles on their path to safety." This source domain supplies agency and intentionality, constructing migrants as purposive actors moving toward a destination against obstacles placed in their way. The entailments differ accordingly. Journeys have travelers, destinations, and impediments, and the impediments here are bureaucratic and legal rather than natural, that is, institutional and therefore alterable.

The two source domains are not positive and negative versions of a single frame. They make different things available for thought: one presents a phenomenon to be contained, the other a course of action to be facilitated or obstructed. Neither, however, constructs migrants as economic actors or as bearers of capacity. This absence is common to all four outlets and is discussed in Section 4.2.

Concordance-based interpretation has known limitations. This is because low-frequency collocations contain limited space for qualitative analysis and because the lines in the concordance contain only part of the context; a full discourse-historical analysis should be done on full articles and their conditions of production. The patterns reported here are thus not based on the power of examples, but rather on their systematicity and their convergence across several independent analytical procedures.

4. Discussion

This section discusses four key findings: first, the ideological split in migration framing (Section 4.1); second, the event-basedness of migration framing (Section 4.2); third, the selective framing of specific migrant groups (Section 4.3); and fourth, the grammatical diffusion of responsibility for migration enforcement (Section 4.4). In the case of each finding, the objective is to identify the mechanism that causes the pattern to occur, to connect it to the theoretical frameworks discussed in Section 1, and to state whether the finding is an extension, confirmation, or qualification of the previous one. The section next discusses the contributions of the study, the implications for policy and practice, and limitations.

4.1. Ideological divergence beneath professional norms

The first has to do with the systematic difference between left- and right-leaning outlets in the use of evaluative polarity, collocation, and metaphor (Tables 6 and 7). This aligns with the threat–humanitarianism axis found in the content analytic literature (Benson, 2013; Chen *et al.*, 2022; Eberl *et al.*, 2018; Liu, 2021). The CL + CDA design can contribute to evidence of the extent to which the divergence is realized. Prior research reveals that outlets vary in the frames that they use; the current analysis reveals that outlets vary in the recurrence of collocation, grammatical voice, and source domain of metaphors. These resources exist beneath the level of frame categories that are usually assessed in content analysis, and also when the manifest framing is balanced. Securitizing presuppositions persist in the use of the editorial voice that Fox News uses to announce a political claim, where the claim itself is being called into question (Section 3.5), and *The Telegraph* securitizes through a managerial vocabulary ("influx," "control," and "security") that does not register as evaluative at all (Section 3.2). Only manifest frames would be coded as neutral in a study, but these are not neutral. This difference described in previous studies is thus replicated here and moved: it has become a characteristic of lexical and grammatical routine and not of declared position.

A comparison with Gabrielatos and Baker (2008) indicates the durability of these routines as well. Their analysis of the UK press between 1996 and 2005 established the conflation of "refugee," "asylum seeker," and "migrant," and the quantification and water metaphors clustering around them. The present corpus documents the same associations two decades later, in a partly different set of outlets, and extends them to a second national system. They survived a change of government, a constitutional realignment, and a transformation of the economic basis of the news industry, and they hold in the US as well as in the UK, which suggests that they are not an artifact of the more openly partisan UK press tradition. The intensification that Gabrielatos and Baker (2008) and KhosraviNik (2010) identified around moments of political contest also takes a more specific form here. It proceeds less through the introduction of new vocabulary than through the densification of existing combinations, as "illegal" and "border" become more frequent alongside node terms whose collocational company was already established. This is consistent with Kelling and Motschenbacher's (2018) findings for German and UK constructions of "illegal migration" and with Taylor's (2021) evidence that migration metaphors persist over long periods even as their evaluation shifts.

Why remain in these patterns, when journalists have jobs to do? Section 1.2 reviews theories that suggest three "reinforcing mechanisms." The first is called naturalization. Fairclough (2015) stresses that ideology is most effective when it is no longer acknowledged as a stance, but is instead "the usual way of talking about something." Van Dijk (2018) also sees the ideological work of the migration discourse in shared and unspoken assumptions, instead of in the actual arguments. This is what this is like in practice, according to the present findings. Securitizing presuppositions exist not only when endorsing a political claim, but also when it is being challenged, as in *The Telegraph*'s language, which purports to be descriptive. Those are the processes not covered by the checks performed in professional practice. A reporter can be neither overtly editorial nor ideologically biased at the level of the rules he or she accepts as governing what he or she does, and nonetheless, be consistently ideological at the collocational level.

Another is institutional reproduction. Unlike conventional CDA, discursive institutionalism provides

an account for what conventional CDA falls short of specifying: the continuity of discursive patterns over changes of personnel and political context (Schmidt, 2008). Institutions perpetuate ideas outside of the intention of any individual member of staff, in style guidebooks, beat structures, source relationships, archives of past coverage accessed when new ones are written, and even audience expectations developed over decades. A journalist writing about a Channel crossing has a vocabulary which is already partially shaped by previous texts on the topic—that is the path of least resistance through which journalists must move. On this evidence, framing will demonstrate more long-lasting durability than any can explain based on the beliefs of current staff, and this mechanism is the best explanation for that 20-year durability. In the same vein are the notions of the "register effect" detected in Section 3.3, which can be explained as due to the influence of institutional norms of register but not to an ideological commitment. This modus vivendi's explicitness would be less exaggerated with the moderated version *The New York Times*, which Hallin and Mancini (2004) describe as "professionalized US journalism" and Hanusch and Hanitzsch (2017) show to be more detached in the US. What *The Telegraph* does find restricting, though, is the broadsheet tradition of voicing restrictionist views as viable policies, even in a more explicitly partisan national newspaper.

The third mechanism is related to the cognitive aspects of collocation and the reasons for its significance beyond the text. Co-occurrence of items can add up to produce semantic prosody, which can then be spread to other instances of the item as described in Section 1.2, and lexical priming creates an expectation which can be triggered on later occurrences of the item without conscious processing (Hoey, 2005). Anyone who has seen a migrant and an illegal immigrant thousands of times over the past 10 years does not need the association to be readily available to him; it already has been acquired through exposure. This could contribute to the understanding of why framing can also change attitudes even among reluctant recipients (Eberl *et al.*, 2018). It is not persuasion, but the piling up of what sounds normal. Thus, on this account, the idea that undocumented immigrant and illegal alien are interchangeable—a distinction of editorial preference—can be thought of as a matter of what each term makes cognitively available.

Two qualifications apply to these conclusions. The patterns described are tendencies rather than uniformities, since Sections 3.4 and 3.5 record securitizing formulations in *The Guardian* and humanitarian formulations in Fox News, and Section 4.2 shows that either can expand considerably under particular conditions. In addition, the claims made here concern textual patterning and its likely cognitive availability, not realized attitudinal effects, which would require reception evidence that this design does not provide.

4.2. Event-driven convergence and the conditionality of ideological framing

The second finding marks the study's main departure from the existing literature: in the event-dominated periods, variation within outlets across the decade was comparable to the variation between outlets within a period (Section 3.4). The literature has generally treated outlet ideology as the principal explanatory factor and event context as a moderator of it (Chen *et al.*, 2022; Eberl *et al.*, 2018; Islentyeva, 2020). The present data do not contradict those studies. However, they question the generality of that ordering. Event context appears to have shaped framing more than outlet alignment did during the Afghanistan evacuation (P3) and the Ukrainian crisis (P4), whereas alignment was the more visible influence in P2 and again in 2023–2024. They cannot be substituted for each other and vary with time. Studies sampling from different windows will thus draw different conclusions regarding the relative importance of ideology, and each may be correct for its own window. This is because of the discourse-historical approach presented in Section 1.2. Following Reisigl and Wodak (2016), there is a need to investigate discursive strategies against the backdrop of the historical and political context in which they are produced, and the results in the present study show that context not only inflects an outlet's stable position but can temporarily alter this position. Emphasizing this view of reading, the ideology provides a set of framing possibilities that events can activate, suppress, and/or redirect. Framing is more conditional than determined. A model which takes editorial alignment as the norm and events as the exceptions will systematically under-describe the moments when it matters most that this is not the norm. Taking the nature of the period subcorpora into account (Section 2.1), these claims are about the relative direction and magnitude of the various factors involved, rather than about their precise relative weights.

In the corpus, three mechanisms of convergence can be discerned. The first is "humanitarian saturation," the idea that by virtue of the moral content of an event, securitism is impossible in an outlet's own moral sensibilities. *Coverage of Afghan* interpreters is of this kind: Fox News grounded its entitlement in service rendered to US forces, striking a humanitarian note without abandoning a merit-based logic of deservingness. The second is frame borrowing under competitive pressure, in which outlets

adopt vocabulary outside their usual repertoire because a story dominating the news agenda cannot otherwise be covered. The third, and analytically the most important, is identity-conditioned convergence, in which the population concerned is positioned so that both ideological repertoires yield the same output. The Ukrainian case is of this third type and is examined in Section 4.3.

Convergence also runs in the opposite direction. *The Guardian*'s use of "small boats crisis" for Channel crossings (Section 3.4) imports a securitizing construction into an outlet with a rights-based baseline profile. The phrase aggregates and problematizes at once. Once a framing circulates widely enough, it becomes the neutral term of reference, simply the name for the thing, and is used even by outlets that would disavow its evaluative entailments if these were made explicit (Parker *et al.*, 2022). This is naturalization in Fairclough's (2015) sense, operating not within but across ideological positions. Wigger *et al.* (2022) report a comparable shift in the liberal German press after the Cologne assaults, and Abbas (2020) describes a parallel policy turn in Denmark and Sweden. The present corpus shows a similar trend. After P1, the UK left-leaning subcorpus contains fewer humanitarian terms and more terms of integration and community cohesion, whereas the US left-leaning subcorpus retains a high proportion of humanitarian terms but, during the Biden administration, acquires more vocabulary of border security and legal pathways.

A further form of convergence lies in what is largely absent from all four outlets. Migrants are constructed as threats or as victims and only rarely as working members of society, whether as taxpayers, colleagues, or bearers of capacity (Section 3.7). The partial exception is *The New York Times*, which does register the economic contribution of undocumented workers. This absence is striking for a period in which the COVID-19 pandemic made migrants' role in essential work unusually visible (Dotsey, 2024), and in which research has documented migrants' economic agency, including their strategic choices as entrepreneurs (Koroutchev, 2025). The victim frame, however well-intentioned, resembles the threat frame in constructing migrants as objects of policy rather than as agents within society (which is why the coding of sympathetic victimhood as positive in Section 2.3 records only the direction of evaluation), and the reduction of people to administrative categories is itself a source of harm (Bhat *et al.*, 2025). The left–right axis around which most of the literature is organized does not register this shared constraint, which may matter more for public understanding than the difference that has attracted most scholarly attention.

4.3. Selective humanitarianism: Racialized and geopolitical deservingness

The third finding concerns the differential treatment of Ukrainian refugees in 2022. Ukrainian refugees attracted humanitarian and solidarity framing across all four outlets, including those whose baseline profiles are clearly securitizing. Fox News, whose collocational profile centers on "invasion," "crime," and "deportation," published lines holding that the US should support Ukrainians fleeing Russian brutality (Fox News, March 2022), and *The Telegraph*, whose profile centers on "influx," "threat," and "control," urged the UK to open its doors (The Telegraph, March 2022). Neither outlet extended comparable framing to Syrian refugees in 2015 or to Central American asylum seekers at any point in the decade; for those populations, "migrant" continued to collocate with "illegal," "invasion," and "crime." Ideology-centered accounts cannot accommodate this pattern. The theories of deservingness reviewed in Section 1.1 can, and the corpus allows three contributing mechanisms to be distinguished.

The first mechanism is racialized proximity. Hierarchies of deservingness in European coverage treat refugees construed as culturally and racially close to receiving publics more sympathetically than those construed as other (Chouliaraki & Georgiou, 2022; Georgiou & Zaborowski, 2017), and Holmes and Castañeda (2016) show that German reporting sorted deserving from undeserving refugees by origin rather than by circumstance. More recent work reaches similar conclusions for public attitudes and official discourse in 2022 (De Coninck, 2023; Garland & Lee, 2025) and for early broadcast coverage of the Ukrainian displacement (Sambaraju & Shrikant, 2023). The present corpus provides Anglophone corroboration and extends these findings in one important respect. Because the differential holds across the ideological spectrum, it operates at a level prior to editorial alignment. Both the securitizing and the humanitarian repertoires appear to carry an unstated restriction on the populations to which they apply. This extension is possible because the corpus varies ideology and population simultaneously, which single-event studies cannot do. It shows that the selectivity documented in the deservingness literature is not a feature of conservative coverage that liberal coverage corrects, but a feature of the repertoire that neither alignment escapes. The categorical politics described by Crawley and Skleparis (2018) and Keyel (2024) explain how this works linguistically: the choice of "refugee" rather than "migrant" is the point at which deservingness is conferred (Sections 3.1 and 3.2), and the corpus shows the same outlets making that choice differently for populations of different origin.

The second mechanism is geo-political alignment. The Ukrainian refugees were coming as the victims of aggression on the part of an enemy government, and their reception might be presented as an act of alliance instead of the accommodation of outside pressure, as an act of welcome instead of surrender. The humanitarian frame also had a material dimension: Ukrainian refugees were offered relatively large access to social support and integration programs in the countries where they landed, while the integration of the Syrian refugees was more dependent on the ambivalent political will and public opinion of their host countries (Amaral *et al.*, 2018). Syrian displacement, occurring in a conflict where Western responsibility was questioned, and Central American migration, a two-way affair between the US and its neighbors, were neither of these situations. Afghan coverage of interpreters is based on this logic, which believes that their entitlement comes from serving a national project, not from a need. Based on this evidence, the use of the humanitarian frame is more an oblique way of using the frame in a right-leaning context, where it does not challenge a nationalist story of national interest, which is an explanation for why seemingly ideological overlaps do not necessarily mean ideological shifts.

The third is a temporal one: It renders the first two mechanisms understandable apart from an overall change of mood. The securitization of small boats arrivals increased in the UK in 2022, and in the same timeframes Ukrainian coverage was present as well as securitizing framing for small boats arrivals and humanitarian framing for arrivals from Ukraine. So convergence can only be explained by more general liberalization in the immigration discourse. It is this simultaneity that is the strongest proof in the corpus that the differential is attached to the population being framed rather than to the period.

These findings are a productive modification of Fairclough's (2015) framework. That model understands ideology as legitimating relations of power between dominant and subordinate groups, through discourse, by making the relations seem natural. The current evidence supports this report at the level of host public or migrant, but the boundary is not as the framework's binary would suggest. It is not hosting publics and migrants as such that divide this corpus, but the one whose claims can be understood in the national or civilizational self-understanding of the host public and the migrants whose claims cannot. The outgroup is not just excluded through ideology, but one outgroup is split, and the split defines the mobilization of a repertoire: either humanitarian or securitizing. Selective humanitarianism is therefore not an aberration in the framing process, but rather one of the mainstays of framing. It can even be a superior means of cutting out many refugees, as the compassion shown to some will be proof against the accusation of prejudice to others.

4.4. Grammatical agency and administrative naturalization

The fourth finding relates to the use of grammatical means of distributing responsibility. They work outside the range where readers and journalists normally perceive framing, and account for why this ideological patterning outlined in Section 4.1 is not detectable by professionals.

The most consequential of these resources is the agentless passive in enforcement reporting. As illustrated in Section 3.6, the institutional actor is often not part of right-leaning narratives of removal, detention, and interdiction. The construction "Migrants were deported" is not merely a stylistic variant of "The Home Office deported migrants"; the former omits the institutional actor responsible for the action. Fairclough (2015) notes that such impersonalization can obscure the human agents behind an action, and the backgrounding or omission of social actors is treated as a nomination strategy in the discourse-historical approach (Reisigl & Wodak, 2016). Read in this light, impersonalization mutes the affective force of enforcement and pre-empts critical engagement with it. In migration reporting, the consequence is that deportation is represented as something that happens to a person rather than as a decision taken by a government.

Two aspects of these constructions are emphasized here. First, they do not stand out when they are alone. In institutional and administrative prose, agentless passives are ordinary—no one would notice that one sentence was tendentious; it is the distribution as a whole that is tendentious, and that can be detected only by corpus methods and not by intuition. Secondly, the constructions emulate the authority of the bureaucratic register. Official communication, recounted in the impersonal voice, makes an implicit claim to procedural neutrality, which is carried over into reporting that uses such a voice and thereby represents challenged enforcement decisions as the normal course of an administration (Unger *et al.*, 2016). This effect could be called not so much "naturalization" as "administrative naturalization"—it is no longer something which needs to be justified, but it is simply removed from the area of justification. It is naturalization in Fairclough's (2015) terms: naturalization through grammar and not vocabulary.

Left-leaning outlets break with this practice by characterizing enforcement actions by identifying the institutional actors involved in the coverage, making the

decisions again political. This is not a neutral alternative, and as a framing it may not be a former object of scrutiny, but the analysis should not be read as a suggestion that foregrounding state agency is the right way to go, as that is itself a frame which has as its objects people and institutions of the government. What matters is that there is a different distribution of responsibility in the two patterns, and that neither of them is a lack of framing.

Metaphorical patterning (Section 3.7) works to the same end. Water metaphors and agentless passives are complementary resources: the former strip migrants of agency, and the latter strip states of it. Combined, as they are in right-leaning coverage, they produce an account in which a natural phenomenon meets an administrative process, and no one acts at all (Charteris-Black, 2011; Lakoff & Johnson, 2003). Johnson *et al.* (2024) show that fluid metaphors carry the same imaginative geography even into migration scholarship, which underlines how deeply conventionalized this source domain has become. The complementarity matters for critical media literacy, because it locates the ideological content of migration reporting outside the evaluative vocabulary where readers are most likely to look for it.

4.5. Contributions to knowledge

Overall, the findings relate to the literature reviewed in Section 1 in three ways. First, they reaffirm the threat–humanitarianism structure that organizes migration coverage in both national systems (Chen *et al.*, 2022; Eberl *et al.*, 2018; Gabrielatos & Baker, 2008; Liu, 2021), and they show that the associations documented for the UK press of 1996–2005 remain valid two decades later and extend to the US. Second, they qualify the ordering on which this literature rests, since at particular moments temporal and event-driven variation can rival ideological variation. Third, they extend it in two respects. The racialized deservingness identified by Chouliaraki and Georgiou (2022) and others cuts across editorial stances rather than being confined to one of them, and ideological divergence is realized less through manifest frame choice than through collocation, grammatical voice, and metaphorical entailment.

These relations translate into three contributions. Empirically, the study provides the first decade-long, same-language comparison of migration framing in UK and US media within an integrated CL + CDA framework, and it finds that the patterns held through major political and industrial change, which favors institutional over intentional explanations of discursive reproduction. Theoretically, it demonstrates that models which focus on outlet ideology as a main causal factor of migration framing are underspecified. Instead, the evidence suggests a multi-dimensional model where tendencies are provided by outlet ideology; pressures are provided by the temporal context; baseline tendencies are shaped by migrant origin and racialization; the potential for reception as a narrative of the national interest is shaped by geopolitical alignment; and the extent to which a stance is expressed is shaped by register. Single-event designs, in which the point of migration origin and the geopolitical alignment of the migrant holders are kept fixed, cannot capture their interplay, and these dimensions are not additive.

On the methodological level, the study demonstrates the usefulness of a longitudinal corpus design and the reflexive assessment which CL + CDA scholarship more and more calls for (Bednarek *et al.*, 2024). The Ukrainian case illustrates the possibilities of the combination of methods. The collocational profiles in Table 6 were constructed using frequency and association measures in the Corpus, and the distinction between *The Telegraph*'s collocational profile and those of Fox News was surfaced by the corpus. This quantitative analysis then identified P4 and 2022, in particular, as an anomaly, which is contrary to what the literature would suggest. The anomaly did not explain itself, though. This meant that concordance reading using knowledge of the literature on deservingness was needed in addition to the temporal comparison discussed in Section 4.3, which is also a corpus affordance. Both quantitative analysis and qualitative analysis were required: quantitative analysis alone would not have identified that this was systematic, and qualitative analysis alone would not have been able to explain this result. The integration consequently does not consist in adding an interpretation to the results of the procedures, but in the cooperation of both procedures where a pattern is detected.

4.6. Implications for policy and practice

The mechanisms identified above have implications for three areas of practice.

First, for journalism, the framing documented here operates not through explicit evaluation but through collocation, grammatical voice, and metaphor, which the profession's existing safeguards do not reach. The patterns reported in Section 3 are compatible with accuracy checks, source balance, and the separation of comment from reporting; the survival of securitizing presuppositions in editorial voice is a case in point. Professional codes and newsroom training would need to address representational practice at the level at which it actually operates: the choice between "undocumented" and "illegal," whether the agent of an enforcement action is named, and whether default metaphors of arrival are adopted. Corpus-based self-audit is a useful tool here, since the relevant patterns are not

discernible in any single editorial decision but are readily measurable in an outlet's output. The aim is not for outlets to abandon editorial positions, which is neither possible nor desirable, but for them to be conscious of how those positions are realized linguistically.

Second, media literacy education has traditionally taught bias detection at the most explicit level of framing, the lexical level. A reader taught to recognize "invasion" as evaluative will not thereby recognize the agentless passive of "migrants were removed," the presupposition of "small boats crisis," or the entailments of a water metaphor. Collocation, grammatical voice, and metaphorical entailment would benefit from explicit treatment in curricula, and this study shows that they are analytically tractable. The register effect has a specific pedagogical implication. *The Telegraph* and *The New York Times* are no less ideologically committed than the outlets that share their positions, but they are less evaluatively explicit; their framing is therefore harder to see and needs to be made visible.

Third, selective humanitarianism poses a problem for media regulation that content-neutral regulation cannot easily resolve. As far as this corpus can show, the differential treatment of Ukrainian and Syrian refugees involved no distortion of fact; what changed was the framing repertoire, and no obvious criterion of accuracy distinguishes the two. This points beyond content toward the structures of the media system: media plurality, public service provision, and the extent to which migrants appear in investigative and long-form journalism as sources and agents rather than only as objects of policy. Threat and victim frames are harder to sustain where migrants speak for themselves, as research in this journal on the political voice of refugees also indicates (de Oliveira, 2023), and the data in Section 3.6 show how rarely migrants are constructed as agents. Finally, because selective humanitarianism is not a matter of hostility, a reduction in overt hostility would not guarantee equitable representation. Addressing only the most visible form of framing would target only its most superficial mechanism.

4.7. Limitations and directions for future research

The study has several limitations, each of which suggests a direction for further research.

First, the four outlets are national and mainstream; tabloid newspapers, broadcast news, local and regional outlets, and digital-native publications are excluded. Because of the register effects identified in Section 3.3, the corpus may understate the range of evaluative patterning in the wider media system. The convergence findings depend least on register, since they concern relative movement across outlets; the agency findings depend on it most, since conventions of grammatical voice vary considerably across registers. Future studies should establish whether the convergence dynamics described in Section 4.2 hold in these other sectors.

Second, the corpus is a thin slice of each outlet's output, averaging 12.5 articles per outlet per year. The design supports period-level comparison, with 25–38 articles per outlet per period, but not year-level or genre-by-year inference, and the diachronic findings in Sections 3.4 and 4.2 should be read as evidence of the direction and approximate scale of change rather than as estimates of annual coverage. Event weighting (Table 1) also means that years of major events are over-represented relative to a proportional sample, which favors the analysis of contested moments over the representation of routine coverage. A larger corpus, ideally the full population of migration-focused articles in each outlet, would permit year-level analysis and robust genre comparison, allow lower-frequency patterns to be tested, and make it possible to check the period-level results reported here against a proportional baseline.

Third, the UK–US design buys analytical control at a cost. It supports no claims about migration framing in non-Anglophone contexts, in the media of the Global South, or in the media of sending countries, where most of the world's migrants are found and where other representational and political frameworks prevail (Juster *et al.*, 2026; Ullah & Chattoraj, 2023; Verma & Kironska, 2024). But the two national subcorpora are also two media systems which are different in terms of their market structure, ownership structure, and audience size; comparability was achieved not just through the common language but also through matched sampling (number of articles, match of the time frame, and match of the topics). Relationships between the UK and the US seen in the context of these four outlets should thus not be viewed as national differences, but rather as outlet differences. An analysis based on the application of the framework to other language media (Spanish, Arabic, German, French, or Mandarin) would indicate whether selective humanitarianism is unique to the discourse in Anglophone receiving countries or a common characteristic of migration reporting, which this study was careful to avoid issues of cross-linguistic comparability (Nartey & Mwinlaaru, 2019).

Fourth, the results are shaped in part by design decisions. Frequency comparisons across the decade are not strictly like-for-like, because digital publishing formats and conventions of article production changed over the period and the boundary between news and comment became less clearly drawn, particularly in the US subcorpus. The

5L–5R window is conventional but not neutral: a narrower window would have favored attributive collocates such as illegal, at the expense of the evaluative-prosody findings, which depend on more distant co-occurrences. Requiring both MI and LL thresholds offsets the opposite biases of the two measures. Lemmatization combines the singular and plural forms of migrant, refugee, and immigrant and may mask a meaningful difference between them, since plural forms aggregate people in a way that singular forms do not. Coding sympathetic victimhood as positive (Section 2.3) records only the direction of evaluation, not its political effect. Other defensible decisions would have produced a somewhat different analysis.

Fifth, not all editorial voice and quotation is clear-cut (Section 3.5). Not all utterances readily lend themselves to classification as direct quotation, indirect reported speech, paraphrase, or free indirect discourse, and outlets do not always indicate whether a source's framing is endorsed. While this uncertainty does not impact the finding that there are some securitizing presuppositions in editorial voice that are based on unambiguous cases, judgments about proportions of quoted and editorial framing require more of a marginal decision.

Sixth, in the Ukraine–Syria comparison, one population is constructed as racially and culturally proximate and the other is not, and one is geopolitically aligned with the receiving states and the other is not. Comparisons of populations for which these two dimensions diverge, for example, racially proximate populations from geopolitically distant states, would allow their relative weight to be estimated. Other dimensions that shape how asylum is adjudicated and understood, such as sexuality and gender identity, have not been explored here and cannot be fully captured by a dichotomy based on national origin (Mole, 2022; Zanin & de Freitas Junior, 2024). Environmental drivers of mobility, such as climate change in sub-Saharan Africa, also produce forms of displacement that the conventional refugee–migrant binary struggles to accommodate (Nzabirinda *et al.*, 2026).

Finally, the analysis concerns production, not reception. Lexical priming, as described in Section 4.1, is not attested in this corpus, but is rather a theoretical framework explored by psycholinguists; the corpus indicates that some connections were accessible to scale readers rather than that readers made them. Anything stated about the effects on attitudes, then, is an expectation that is built on a theory. Experimental designs that manipulate collocation, grammatical voice, and metaphorical frame in a controlled way such that the propositional content of the message remains the same would help to determine whether, when not directly evaluated, attitudes change in relation to these factors, and would bring the quantitative corpus-based discourse analysis more closely into line with research on the effects of discourse in the media.

5. Conclusion

This study set out to characterize the ideological framing of migration in UK and US national news media and to determine how consistent that framing remained over ten years of political change. Its answer is that ideology matters, but not in the way the field has typically assumed. Outlet alignment supplies durable linguistic routines that persist largely below the level of awareness, but these routines are conditional: events can suspend them, and the identity of the population represented determines which of them is applied. The central point is therefore not that left- and right-leaning outlets differ, which is well known, but that both draw on a shared set of resources that distribute compassion selectively and conceal responsibility for enforcement in grammar. The selectivity begins with the label itself, since the choice between "refugee" and "migrant," and between "undocumented" and "illegal" as modifiers of "immigrant," can confer or withhold legitimacy before any argument is made.

The study suggests that single-event and cross-sectional designs should be interpreted with caution, because the apparent weight of ideology depends on the period sampled, and that analyses should attend as closely to who is represented as to who is doing the representing. For journalism and media literacy, it calls for attention to shift from overt bias to the subtler work of collocation, voice, and metaphor through which migration is routinely discussed. Making these choices visible is a first step toward fairer representation of migrants and refugees in public discourse.

Acknowledgements

None.

Funding

None.

Conflict of interest

The authors declare that they have no financial or personal relationship that may have inappropriately influenced them in writing this article.

Author contributions

Conceptualization: All authors
Formal analysis: All authors

Methodology: All authors
Writing – original draft: All authors
Writing – review & editing: All authors

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data

The corpus consists of publicly available news articles; no human participants were involved. Because the full texts are subject to copyright, the corpus cannot be shared publicly. The article list (outlet, date, headline, and URL), the coding manual, and the frequency and collocation outputs are available from the corresponding author upon reasonable request.

AI tools statement

None.

References

Abbas, H. (2020). *Immigration and integration policy in Europe: Denmark and Sweden, 1970–2010*. Taylor & Francis.

Amaral, E. F. L., Woldetsadik, M. A., & Armenta, G. (2018). Challenges to the integration of Syrian refugees. *International Journal of Population Studies, 4*(1), 39–56.

https://doi.org/10.18063/ijps.v4i1.457

Anthony, L. (2022). *AntConc* (Version 4.2.0) [Computer software]. Waseda University. Accessed January 24, 2024. https://www.laurenceanthony.net/software

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh University Press.

Baker, P. (2023). *Using corpora in discourse analysis* (2nd ed.). Bloomsbury Academic.

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society, 19*(3), 273–306.

https://doi.org/10.1177/0957926508088962

Bednarek, M., Schweinberger, M., & Lee, K. K. H. (2024). Corpus-based discourse analysis: From meta-reflection to accountability. *Corpus Linguistics and Linguistic Theory*, *20*(3), 539–566.

https://doi.org/10.1515/cllt-2023-0104

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Benson, R. (2013). *Shaping immigration news: A French-American comparison*. Cambridge University Press.

Bhat, N., Paul, F. A., Bashir, A., Gul, A., & Ganie, Z. A. (2025). A narrative review of cultural identity and health risks in forced migration. *International Journal of Population Studies, 11*(2), 1–16.

https://doi.org/10.36922/ijps.4468

Brezina, V. (2018). Statistical choices in corpus-based discourse analysis. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review*. Routledge.

Brezina, V., Weill-Tessier, P., & McEnery, A. (2021). *#LancsBox* (Version 6.x) [Computer software]. Lancaster University. Accessed February 17, 2024. http://corpora.lancs.ac.uk/lancsbox

Charteris-Black, J. (2011). *Politicians and rhetoric: The persuasive power of metaphor*. Palgrave Macmillan.

Chavez, L. R. (2021). *The Latino threat: Constructing immigrants, citizens, and the nation* (3rd ed.). Stanford University Press.

Chen, K., Babaeianjelodar, M., Shi, Y., Janmohamed, K., Sarkar, R., Weber, I., Davidson, T., De Choudhury, M., Huang, J., Yadav, S., Khudabukhsh, A., Nakov, P. I., Bauch, C., Papakyriakopoulos, O., Khoshnood, K., & Kumar, N. (2022). *Partisan US news media representations of Syrian refugees* [Preprint]. *arXiv*. Accessed January 27, 2024. https://arxiv.org/abs/2206.09024

Chouliaraki, L., & Georgiou, M. (2022). *The digital border: Migration, technology, power*. New York University Press.

Colbern, A., Walker, S., Glenn, K. M., Schmidt, R., & Harrigan, J. (2024). Crisis framing of immigration politics in United States newspapers, 1980–2022. In S. M. Tsegay (Ed.), *Refugees and migrants: Current conditions and future trends* (pp. 1–25). IntechOpen.

https://doi.org/10.5772/intechopen.1005525

Consterdine, E., & Hampshire, J. (2014). Immigration policy under New Labour: Exploring a critical juncture. *British Politics, 9*(3), 275–296.

https://doi.org/10.1057/bp.2013.19

Crawley, H., McMahon, S., & Jones, K. (2016). *Victims or villains? Perceptions of migrants in the UK*. Overseas Development Institute.

Crawley, H., & Skleparis, D. (2018). Refugees, migrants, neither, both: Categorical fetishism and the politics of bounding in Europe's 'migration crisis'. *Journal of Ethnic and Migration Studies, 44*(1), 48–64.

https://doi.org/10.1080/1369183X.2017.1348224

De Coninck, D. (2023). The refugee paradox during wartime in Europe: How Ukrainian and Afghan refugees are (not) alike. *International Migration Review, 57*(2), 578–586.

https://doi.org/10.1177/01979183221116874

De Genova, N. (2013). Spectacles of migrant 'illegality': The scene of exclusion, the obscene of inclusion. *Ethnic and Racial Studies, 36*(7), 1180–1198.

https://doi.org/10.1080/01419870.2013.783710

de Oliveira, L. G. (2023). Refugees and refugee applicants' voices in Brazilian cities. *International Journal of Population Studies, 9*(3), 84–97.

https://doi.org/10.36922/ijps.438

Del Fante, A. (2018). Where are 'they' from? A corpus-assisted study of the geographical representation of migrants in the UK and Italian quality press in 2016 and 2017. *I-LanD Journal: Identity, Language and Diversity, 1*, 28–52.

https://doi.org/10.26379/IL2018001_003

Del Fante, D. (2024). *Migrating across times and cultures: Metaphorical images of migration in the U.S. and Italian newspaper discourse between the 20th and 21st centuries*. Peter Lang.

https://doi.org/10.3726/b21857

Dotsey, S. (2024). COVID-19 and the precarious low-skilled workforce in the European Union: Time to call the shots? *International Journal of Population Studies, 10*(3), 114–120.

https://doi.org/10.36922/ijps.2165

Eberl, J.-M., Meltzer, C. E., Heidenreich, T., Herrero, B., Theorin, N., Lind, F., Berganza, R., Boomgaarden, H. G., Schemer, C., & Strömbäck, J. (2018). The European media discourse on immigration and its effects: A literature review. *Annals of the International Communication Association, 42*(3), 207–223.

https://doi.org/10.1080/23808985.2018.1497452

Evert, S. (2008). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (Vol. 2, pp. 1212–1248). Mouton de Gruyter.

Fairclough, N. (2015). *Language and power* (3rd ed.). Routledge.

Financial Times. (2025, May 18). Labour's immigration crackdown. Accessed October 19, 2025. https://www.ft.com/content/8579af75-0d05-4ab6-abe8-4792620f41c0

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning, 67*(S1), 155–179.

https://doi.org/10.1111/lang.12225

Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 225–258). Routledge.

Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996–2005. *Journal of English Linguistics, 36*(1), 5–38.

https://doi.org/10.1177/0075424207311247

Garland, J., & Lee, J. (2025). Ukrainian refugees and welfare deservingness: A comparative study of UK government discussions around the 2022 Ukraine conflict and 2015 migrant crisis. *The British Journal of Sociology, 76*(4), 800–813.

https://doi.org/10.1111/1468-4446.13219

Georgiou, M., & Zaborowski, R. (2017). *Media coverage of the 'refugee crisis': A cross-European perspective*. Council of Europe.

Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge University Press.

Hallin, D. C., & Mancini, P. (2017). Ten years after *Comparing media systems*: What have we learned? *Political Communication, 34*(2), 155–171.

https://doi.org/10.1080/10584609.2016.1233158

Hanusch, F., & Hanitzsch, T. (2017). Journalistic roles and everyday work. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (pp. 167–181). Routledge.

Hart, C. (2020). *Researching discourse: A student guide*. Routledge.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.

Holmes, S. M., & Castañeda, H. (2016). Representing the 'European refugee crisis' in Germany and beyond: Deservingness and difference, life and death. *American Ethnologist, 43*(1), 12–24.

https://doi.org/10.1111/amet.12259

Hu, H., & Cai, C. (2025). Representations of left-behind children in Chinese English-language news media: A corpus-assisted discourse analysis. *International Journal of Social Science, Management and Economics Research, 3*(2), 49–70.

https://doi.org/10.61421/IJSSMER.2025.3204

Hunston, S. (2022). *Corpora in applied linguistics* (2nd ed.). Cambridge University Press.

International Organization for Migration. (2019). *Glossary on migration* (International Migration Law No. 34).

Islentyeva, A. (2020). *Corpus-based analysis of ideological bias: Migration in the British press*. Routledge.

https://doi.org/10.4324/9780429263064

Johnson, L., Jacobsen, M. H., & Ehrkamp, P. (2024). The work of fluid metaphors in migration research: Geographical imaginations and the politics of writing. *Progress in Human Geography, 48*(6), 843–860.

https://doi.org/10.1177/03091325241280398

Jurkowitz, M., Mitchell, A., Shearer, E., & Walker, M. (2020). *U.S.*

media polarization and the 2020 election: A nation divided. Pew Research Center. Accessed November 17, 2025. https://www.pewresearch.org/journalism/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided/

Juster, S. H., Munsell, J. F., Njenga, M., & Okia, C. A. (2026). Tree-based intervention typologies and improvements in refugee displacement regions. *International Journal of Population Studies, 12*(1), 15–31.

https://doi.org/10.36922/IJPS025220105

Kelling, K., & Motschenbacher, H. (2018). The discourse of 'illegal migration' in German and British newspapers. *Critical Discourse Studies, 15*(5), 450–466.

Keyel, J. (2024). Immigrants, asylum-seekers, and refugees: Navigating conceptual challenges through multidimensional migration space-time. *International Journal of Population Studies, 10*(4), 40–44.

https://doi.org/10.36922/ijps.3004

KhosraviNik, M. (2010). The representation of refugees, asylum seekers and immigrants in British newspapers: A critical discourse analysis. *Journal of Language and Politics, 9*(1), 1–28.

https://doi.org/10.1075/jlp.9.1.01kho

KhosraviNik, M. (2017). Right-wing populism in the West: social media discourse and echo chambers. *Insight Turkey, 19*(3), 53–68.

https://doi.org/10.25253/99.2017193.04

Koroutchev, R. (2025). Choice of a country for migrant entrepreneurship. *International Journal of Population Studies, 11*(2), 118–131.

https://doi.org/10.36922/ijps.4447

Krzyżanowski, M., Triandafyllidou, A., & Wodak, R. (2018). The mediatization and the politicization of the 'refugee crisis' in Europe. *Journal of Immigrant & Refugee Studies, 16*(1–2), 1–14.

https://doi.org/10.1080/15562948.2017.1353189

Kwak, H., An, J., & Ahn, Y.-Y. (2020). *A systematic media frame analysis of 1.5 million New York Times articles from 2000 to 2017* [Preprint]. *arXiv*.

https://doi.org/10.48550/arXiv.2005.01803

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. University of Chicago Press. (Original work published 1980)

Liu, S. J. S. (2021). Framing immigration: A content analysis of newspapers in Hong Kong, Taiwan, the United Kingdom, and the United States. *Politics, Groups, and Identities, 9*(4), 759–783.

https://doi.org/10.1080/21565503.2019.1674162

Liu, Y. (2024). A corpus-based critical discourse analysis of news reports on the 2023 Israel–Hamas war. *Journal of Linguistics and Communication Studies, 3*(3), 70–84.

https://doi.org/10.56397/JLCS.2024.09.09

Machin, D., & Mayr, A. (2012). *How to do critical discourse analysis: A multimodal introduction*. Sage.

Markowitz, D. M., & Slovic, P. (2020). Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences, 117*(17), 9260–9269.

https://doi.org/10.1073/pnas.1921790117

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

Mayblin, L., Davies, T., Isakjee, A., Turner, J., & Yemane, T. (2024). Small boats, big contracts: Extracting value from the UK's post-Brexit asylum 'crisis'. *The Political Quarterly, 95*(2), 253–262.

https://doi.org/10.1111/1467-923X.13412

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Mole, R. C. M. (2022). *Queer migration and asylum in Europe*. UCL Press.

Nairi, A. H. (2025). The representation of sub-Saharan African migrants in the Tunisian written media: A corpus-assisted critical discourse analysis. *The Journal of North African Studies, 30*(2), 259–293.

https://doi.org/10.1080/13629387.2024.2396333

Nartey, M., & Mwinlaaru, I. N. (2019). Towards a decade of synergising corpus linguistics and critical discourse analysis: A meta-analysis. *Corpora, 14*(3), 391–418.

https://doi.org/10.3366/cor.2019.0169

Nzabirinda, E., Nkurunziza, J., & Mugenzi, M. (2026). Being able to breathe again: Impact of climate change on human mobility in sub-Saharan Africa. *International Journal of Population Studies, 12*(3).

https://doi.org/10.36922/IJPS025210098

Parker, S., Bennett, S., Cobden, C. M., & Earnshaw, D. (2022). 'It's time we invested in stronger borders': Media representations of refugees crossing the English Channel by boat. *Critical Discourse Studies, 19*(4), 348–363.

https://doi.org/10.1080/17405904.2021.1920998

Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. John Benjamins.

Phelan, S. (2022). *Media, journalism and ideology*. Routledge.

Rayson, P., Berridge, D., & Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. In G. Purnelle, C. Fairon, & A. Dister (Eds.), *Le poids des mots: Proceedings of the 7th International Conference on Statistical Analysis of Textual Data (JADT*

2004) (Vol. 2, pp. 926–936). Presses universitaires de Louvain.

Reisigl, M., & Wodak, R. (2016). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse studies* (3rd ed., pp. 23–61). Sage.

Sambaraju, R., & Shrikant, N. (2023). 'All of a sudden for no reason they've been displaced': Constructing the 'contingent refugee' in early media reports on the Ukrainian refugees. *British Journal of Social Psychology, 62*(4), 1715–1732.

https://doi.org/10.1111/bjso.12652

Schmidt, V. A. (2008). Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science, 11*, 303–326.

https://doi.org/10.1146/annurev.polisci.11.060606.135342

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins.

Stefanowitsch, A. (2006). Words and their metaphors: A corpus-based approach. In A. Stefanowitsch & S. Th. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 63–105). Mouton de Gruyter.

Taylor, C. (2021). Metaphors of migration over time. *Discourse & Society, 32*(4), 463–481.

https://doi.org/10.1177/0957926521992156

Ullah, A. A., & Chattoraj, D. (2023). Dispossession and displacement: The crisis and media influence. In *The unheard stories of the Rohingyas* (pp. 63–106). Bristol University Press.

Unger, J. W., Wodak, R., & KhosraviNik, M. (2016). Critical discourse studies and social media data. In D. Silverman (Ed.), *Qualitative research* (4th ed., pp. 277–293). SAGE.

van Dijk, T. A. (2018). Discourse and migration. In R. Zapata-Barrero & E. Yalaz (Eds.), *Qualitative research in European migration studies* (pp. 227–245). Springer.

Verma, M., & Kironska, K. (2024). Exploring public opinions on Rohingya refugees residing in India: An empirical study. *International Journal of Population Studies, 10*(4), 45–62.

https://doi.org/10.36922/ijps.2174

Wigger, I., Yendell, A., & Herbert, D. (2022). The end of 'Welcome Culture'? How the Cologne assaults reframed Germany's immigration discourse. *European Journal of Communication, 37*(1), 21–47.

https://doi.org/10.1177/02673231211012173

Wodak, R., & Meyer, M. (Eds.). (2016). *Methods of critical discourse studies* (3rd ed.). Sage.

Yu, Y., & Brookes, G. (2024). COVID-19 news in China and the UK: A corpus-based critical discourse analysis of *China Daily* and *The Times*. *Corpora, 19*(3), 285–315.

https://doi.org/10.3366/cor.2024.0315

Zanin, H., & de Freitas Junior, A. R. (2024). LGBTQI+ asylum seekers and refugees: Recognition and public policies for protection. *International Journal of Population Studies, 10*(2), 103–114.

https://doi.org/10.36922/ijps.0871