

Mapping the semantic transformations of major powers in Cold War East Asia: a word embedding approach using Chinese and Korean newspapers (1946–95)

Byungjun Kim¹, Jonghyun Jee^{2,3}, and Donghyun Woo^{4,*}

¹Cultural Informatics, Graduate School of Korean Studies, The Academy of Korean Studies, Seongnam, 13455, Republic of Korea

²Department of Communication Studies, Northwestern University, Evanston, Illinois, 60208, USA

³Center for Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology, Daejeon, 34141, Republic of Korea

⁴School of Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology, Daejeon, 34141, Republic of Korea

*Corresponding author. 291 Daehak-ro, School of Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejeon, 34141, Republic of Korea. E-mail: dhwoo1234@kaist.ac.kr

Abstract

Leveraging dynamic word embedding techniques, this study examines semantic shifts surrounding political entities during and after the Cold War as portrayed in influential Chinese and Korean newspapers, and probes how these shifts reveal changing geopolitical perceptions of major powers in these East Asian societies. Our corpus comprises 2.7 million articles from two major East Asian newspapers spanning the period from 1946 to 1995: the Chinese-language *Renmin Ribao* and the Korean-language *Chosun Ilbo*. By tracking the semantically relevant words surrounding the term ‘Soviet Union’ in both dailies, we show how these publications informed the evolving diplomatic and ideological perceptions of the nuclear-armed Cold War superpower. This study demonstrates how a dynamic word embedding-driven approach complements close reading by providing hermeneutic contexts for textual analysis at scale, illuminating long-term perspectives through which Chinese and Korean societies perceived powers and their geopolitical stakes in the Cold War world.

Keywords dynamic word embedding; semantic shifts; Chosun Ilbo; Renmin Ribao; Cold War.

1. Introduction

Among print media, newspapers were central in shaping public understanding of ideological rivalries and strategic alignments during the global Cold War (Westad 2005; Fainberg 2021). Although historical scholarship increasingly addresses East Asia (Li 2024; Tinn 2025), comparative studies of daily newspapers across linguistic and ideological boundaries in the region remain scarce, largely due to uneven digitization and limited access to sources. Yet such comparisons are crucial for exploring how states used domestic media to engage with the

Cold War, revealing shared motives and divergent strategies in media discourse. This approach also broadens the study of media and statecraft beyond its conventional Euro-American focus. In this sense, a comparative analysis of the People’s Republic of China (PRC) and the Republic of Korea (ROK, South Korea) offers valuable insights. Both achieved economic and technological development under authoritarian regimes, albeit with differing political systems. In each case, the media, whether through direct state control or indirect state-media relationships, played a central role in shaping public opinion to support statecraft and diplomacy.

Examining how these societies were guided to perceive foreign powers through the press (Lenoe 2004) helps illuminate the domestic foundations of modern East Asian geopolitics.

Both the PRC and ROK utilized newspapers to inform the public about issues spanning from the local to the global. In the PRC, *Renmin Ribao* (RMRB) has served as the primary mouthpiece for the party-state's ideological directives, while in the ROK, *Chosun Ilbo* (CSIB) has sustained the nation's largest subscriber base (Kim 2018). Examining these newspapers in tandem provides a valuable lens through which researchers can computationally reconstruct how Chinese and Korean societies were guided to view, interpret, and experience the unfolding dynamics of the Cold War.

Despite East Asia's critical role in the Cold War (Masuda 2015), existing scholarship remains predominantly grounded in Anglophone perspectives, with limited attention to the region's interconnected histories. Even among historians focusing on the ideological East (Radchenko 2009; Jersild 2014; Friedman 2015), the emphasis has largely fallen on high-level bureaucratic and diplomatic interactions, relegating media representations to the realm of propaganda. As a result, little attention has been paid to the local narratives of global shifts. The field's relative disengagement from computational methodologies has further reinforced this tendency, leaving selective close reading as the dominant mode of analysis. This has left gaps in our understanding of how portrayal of Cold War powers was deployed in the region's press. For example, although the Sino-Soviet relations deteriorated in the 1960s, few studies have examined how this geopolitical shift was represented in Chinese media, let alone how such portrayals compared with those in South Korea.

This study seeks to address these gaps by leveraging computational analytics to examine the evolving representations of major Cold War powers in the influential newspapers of both the PRC-led Sinophone sphere and the ROK-led Korean-language readership. Applying Dynamic Word Embedding (DWE) techniques to a corpus of 2.7 million articles from RMRB and CSIB spanning 1946–95, we trace shifting semantic contexts of the Soviet Union as a primary case study. Our focus is motivated by two factors. First, preliminary analyses indicate notable shifts in how the Soviet Union was represented in both newspapers. Second, as the leading socialist state, the Soviet Union occupied a central position in Cold War geopolitics. Its portrayal in the press thus offers a crucial lens through which to examine patterns of alliance, rivalry, and ideological tension that shaped domestic discourse.

We demonstrate how word embedding techniques offer a scalable approach to analyze large textual corpora, uncovering insights that would be difficult to obtain solely through qualitative approaches. By capturing temporal shifts in the semantic contexts surrounding specific keywords, these techniques enable a deeper interpretive engagement with historical texts that might otherwise remain inaccessible due to their scale or limited digital availability. Our study contributes to the ongoing discussion about bridging data-driven analysis with the hermeneutic practices of humanities and social science research (Guldi 2024). Moreover, our findings provide a productive framework for evaluating scholarly narratives concerning the societal circulation and reception of Cold War developments. By grounding our interpretive claims in digital evidence, we explore how Chinese and Korean societies' perceptions of major powers and the evolving dynamics of Cold War geopolitics were shaped.

2. Theoretical and methodological background

2.1 Tracking semantic changes

Words and their meanings and usages change over time. This linguistic phenomenon, conceptualized as semantic change, takes place as words are used in varying contexts. The study of semantic change was rooted in historical linguistics, with seminal contributions by Leonard Bloomfield's *Language* (1933) providing the foundational framework for understanding the diachronic evolution of word meanings and their categorization. Later, semantic change research built on Firth's (1957) articulation of the distributional hypothesis, which posited that words appearing in similar contexts tend to have related meanings. While this idea had been difficult to verify due to technological limitations of the time, it eventually became a foundation for computational semantic analysis.

The advent of large-scale digitized corpora and advances in computational techniques, namely Natural Language Processing (NLP), enabled algorithmic tracking of semantic changes (Traugott 2017). This approach assumed that changes in a word's collocational patterns reflected changes in its meaning (Hilpert 2008). A notable development in NLP was the introduction of word embeddings, which represented words as dense vectors in a continuous vector space (Jurafsky and Martin 2000). Sagi et al. (2009) were among the first to propose comparing the density of semantic vector clusters to make statistical inferences about whether a word's meaning had changed across time. A breakthrough in data availability came with Google Books corpora, which digitized

approximately four percent of all books ever printed (Michel et al., 2011). Building upon the Google Books data, Gulordava and Baroni (2011) and Wijaya and Yeniterzi (2011) demonstrated the efficacy of distributional semantic models in detecting semantic change within large diachronic corpora.

Static word embeddings, such as Word2Vec (Mikolov et al., 2013) and GloVe (Pennington et al., 2014), have been widely used in NLP for analytical tasks, including text classification and sentiment analysis. But they perform less successfully in capturing how word meanings change over time, a key limitation for historical language analysis. These circumstances prompted the development of DWE (Bamler and Mandt 2017), which is central to our approach. By capturing how word meanings changed over time, DWE allows the analysis of mid- or long-term semantic shifts in historical texts. DWE extends the concept of word representations to trace how the meaning and usage of words evolve across different periods. These methods have been proven suitable for capturing semantic changes in large general corpora (Hamilton et al., 2016; Kutuzov et al., 2018; Tahmasebi et al., 2021). Recently, word embeddings have been used to examine semantic shifts in historical contexts—shifts in gender bias in historical newspapers (Wevers 2019), changes in gender and ethnic stereotypes (Garg et al., 2018), and the framing of immigrants in political speeches (Card et al., 2022).

2.2 Newspapers in digital humanities research

Nicholson (2013) noted that digital newspaper archives could inspire new approaches to cultural and transnational history. Since then, the digitization of historical newspapers has revolutionized the field, with scholars adopting computational methods to analyze large-scale corpora. Yet, Wijfjes (2017) pointed out that few historical studies using digital sources have undergone thorough evaluation. Newspapers serve as a valuable resource for analyzing societal changes over time as they exist both as historical records and reflections of contemporary culture. They capture contemporaneous societal attitudes, public discourse, and institutional perspectives, offering windows into historical events, policy developments, and cultural trends. Partially structured yet diverse in content, from editorials to advertisements, newspapers offer a rich dataset for computational analysis.

Scholars have used computational methods to analyze newspaper corpora. Ghasiya and Okamura (2021) applied topic modeling and sentiment analysis to examine Middle East-related articles from Japanese

newspapers, suggesting Japan's reliance on Middle Eastern stability for energy resources. Similarly, Jiang (2024) employed word co-occurrence networks and high-frequency emotion words to compare the framing of the Russia–Ukraine conflict in British (Reuters) and Chinese (Xinhua) media. Both studies are relatively narrow in scope, as are their datasets: Ghasiya and Okamura drew on 1,672 articles from three newspapers spanning 2013–20, while Jiang examined 8,859 articles from two newspapers covering 2022–23. Although these studies captured short-term trends in media representation, their limited temporal and thematic scope prevented them from revealing long-term semantic shifts.

Major newspapers provide ideal datasets for understanding how media representation, language, and societal values evolved. For instance, Costa (2018) analyzed Portuguese and Brazilian newspapers from the 1990s to explore different portrayals of Japan. Likewise, Pang et al. (2024) combined topic modeling and critical discourse analysis to investigate representations of Chinese migrant workers in Chinese newspapers. While these studies testified to the value of newspapers for comparative historical research, their methods did not account for the fluidity of semantics over time.

DWE addresses some of these limitations by quantifying semantic shifts. Hengchen et al. (2021) pioneered this approach by leveraging neural word embeddings to study semantic shifts in the concept of nationhood across Dutch, Swedish, Finnish, and English newspaper archives. Particularly relevant to our study is the work of Chen et al. (2024), which employed similar techniques to track the moralization of effort versus efficiency across Chinese and American political corpora, including the RMRB. Their work showed how diachronic analysis of large-scale textual data could reveal evolving cultural attitudes embedded in official discourse, indicating the potential for cross-lingual diachronic analysis in the field of history.

3. Methodology

This section describes the data corpus, the preprocessing steps undertaken to refine the data, and the DWE method used to track semantic changes across time periods. Figure 1 provides a visual overview of the model and data processing pipeline.¹

3.1 Data overview

Our study uses a dataset drawn from the digitized archives of two influential East Asian newspapers: RMRB from the PRC and CSIB from the ROK. Both dailies have

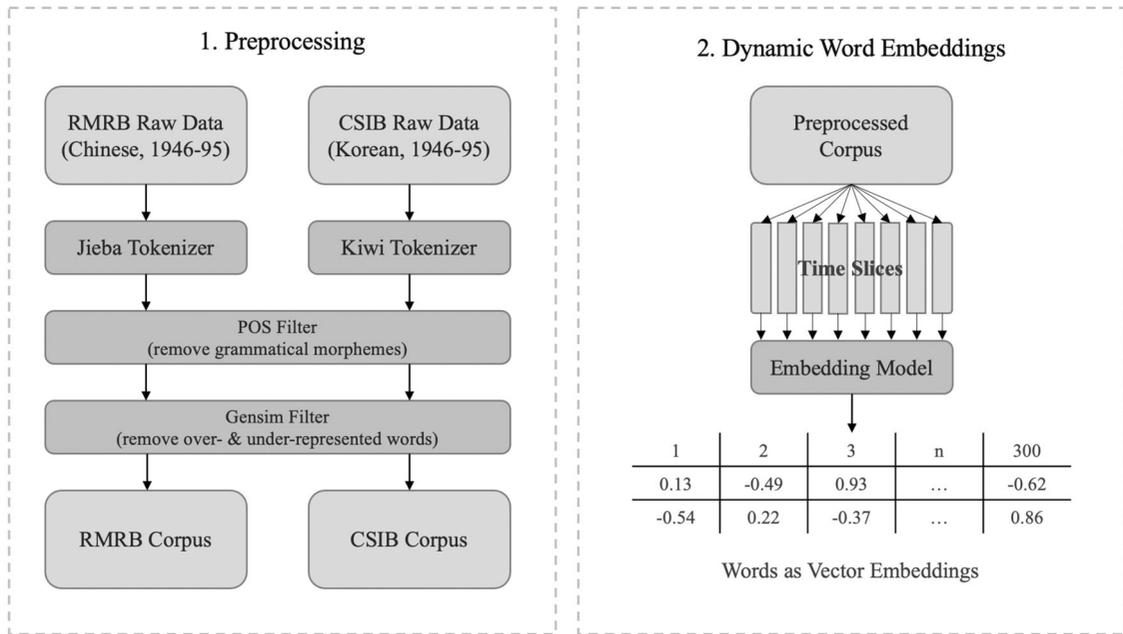


Figure 1. Overview of the processing pipeline.

played a key role in shaping public discourse within their respective societies and broader diaspora communities. While RMRB, established in 1946, has functioned as the official mouthpiece of the Chinese Communist Party (CCP; Yi 2013), CSIB has served as a leading media outlet in the ROK, yielding substantial influence on public opinion since its foundation in 1920 (Han and Lee 2013).

We obtained the RMRB dataset from the Harvard Library and the CSIB dataset from the Chosun News Library website.² The combined corpus spans the period from 1946 to 1995, covering the Cold War era (mid-1940s–91) and its immediate aftermath. Our final dataset comprises 1,039,002 RMRB articles (568,070,216 tokens) and 1,652,116 CSIB articles (515,579,128 tokens) (Fig. 2). To maintain data integrity, we focused on regular news articles from CSIB, excluding advertisements, serial novels, and other non-news content.

3.2 Data preprocessing

The data preprocessing involved several steps to prepare the textual data for effective analysis. Initially, tokenization, a prerequisite in NLP, was performed to break down article texts into smaller elements suitable for further processing. For the RMRB dataset, we employed Jieba³ that offers part-of-speech (POS) tagging for toponyms, as it has demonstrated consistent performance across various types of modern Chinese corpora. For the CSIB dataset, we utilized Kiwipiemy (Lee 2024), a state-

of-the-art Korean morphological analyzer known for its speed and accuracy. This step produced tokens, each tagged with its corresponding POS, such as noun, verb, or adjective.

Not all tokens are equally informative for semantic analysis. Tokens, especially those with grammatical functions, hardly contribute to the meaning of the text. We therefore filtered the tokenized data to retain only those morphemes most relevant to understanding semantic shifts. We excluded tokens based on their POS tags to ensure that our analysis focused on content-bearing words (see Table 1).

We also considered orthographic variation, particularly alternative spellings of key terms. An important case from the CSIB corpus is the spelling of the Soviet Union. Across the dataset, the standard form ‘소련’ appears 88,213 times, while the variant ‘쏘련’ occurs 368 times. In the 1946–50 window that holds the highest frequency of the variant, ‘쏘련’ appears 5,364 times compared to 193 instances of ‘소련.’ The variant declines rapidly afterward (e.g. 115 instances in 1951–5, 30 in 1956–60, and single digits thereafter). Given this disparity, we retained both forms but did not normalize the variant, as its low frequency is unlikely to affect the DWE model.

This consistency also reflects the structure of the CSIB corpus, which includes both original texts (written in mixed Korean and Sino-Korean scripts) and their modern transliterations. Since transliterations are only available

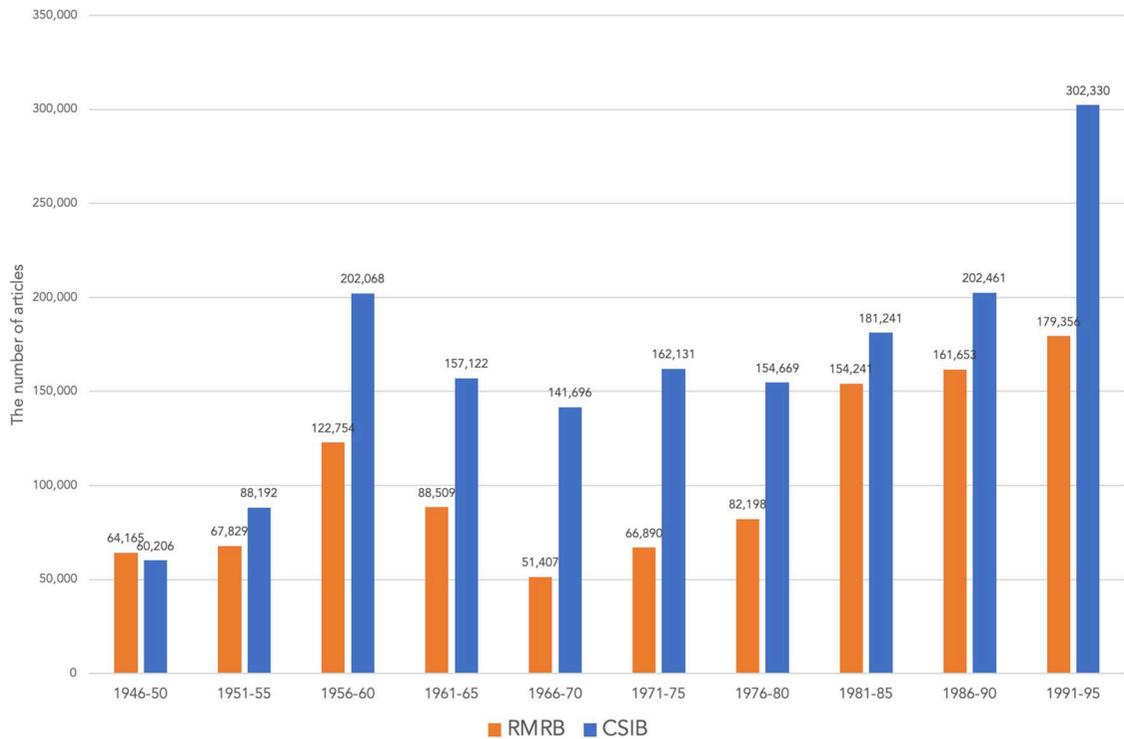


Figure 2. Comparison of article counts between RMRB and CSIB from 1946 to 1995.

Table 1. The summary of the tokenization process.

Dataset	Excluded POS tags	The number of filtered unique tokens
RMRB	‘vd’ (动词副词): adverbial verb ‘ad’ (副形容词): adverbial adjective ‘d’ (副词): adverb ‘m’ (数量词): numeral ‘q’ (量词): classifier/measure word ‘r’ (代词): pronoun ‘p’ (介词): preposition ‘c’ (连词): conjunction ‘u’ (助词): particle ‘xc’ (其他虚词): other function words ‘w’ (标点符号): punctuation ‘x’ (其他): other	1,492,753
CSIB	‘E’ (어미): endings ‘J’ (조사): particles ‘S’ (외국어): foreign words ‘W’ (기타): others	1,194,815

up to 1953, we used the transliterated versions up to that year and the original texts from 1954 onward. Despite this transition, the corpus maintains orthographic consistency throughout.

We then selected specific words for inclusion in the DWE training process. Rather than using the entire corpus, which would cause inefficiency and skew the results, we deployed frequency thresholds to minimize noise and ensure the robustness of research outcomes. To set thresholds, we introduced the parameters of the Gensim⁴ package’s ‘filter_extremes’ function. While the ‘no_below’ parameter sets a lower boundary, excluding words that appear in fewer than a specified number of documents, the ‘no_above’ parameter gives an upper limit, removing words that appear in more than a certain percentage of documents (see Table 2 for specific values). This filtering reduces the overrepresentation of less significant terms.

3.3 Dynamic word embedding

In implementing DWE, we adopted a five-year time unit, generating ten time slices from 1946–50 to 1991–5. This window balances temporal granularity with model stability: ten-year spans risk obscuring media responses to key geopolitical events, while shorter intervals, such as

Table 2. The word selection for DWE training.

Dataset	No_below	No_above	The number of unique tokens for DWE
RMRB	5,000	0.9	4,791
CSIB	10,000	0.9	2,432

one or two years, often yield insufficient data and exaggerate transient lexical shifts. The five-year span mitigates such volatility, enabling more robust modeling of linguistic change over time.

We also configured key hyperparameters that were essential for robust DWE modeling. The dimensionality of the word embeddings was set to 300, which provides a strong balance between context capture and efficiency. Previous research shows that beyond 300 dimensions, performance seldom improves (Patel and Bhattacharyya 2017). The window size was set to three, meaning that three words before and after the target term were considered during the embedding process. While typical word embedding processes use a window size between three and five (Di Gennaro et al., 2021), we chose three to prioritize capturing the closest contextual relationships.

We employed Dynamic Bernoulli Embeddings (DBE; Rudolph and Blei 2018) due to their frequent use in research and the availability of a PyTorch implementation.⁵ Unlike models that generate separate embeddings for each time slice, DBE trains a unified model across all periods, preserving a consistent vector space that facilitates direct comparison of word meanings over time. This makes DBE distinctively effective for tracing diachronic semantic shifts. The model supports three key functions: (1) measuring ‘absolute drift,’ or the degree of semantic change between the first and last time slice; (2) identifying ‘embedding neighborhoods,’ or semantically similar words within a time slice; and (3) detecting change points, which quantify the magnitude of semantic shift between consecutive time periods. These three functions work complementarily: absolute drift identifies words worthy of investigation, embedding neighborhoods reveal their evolving semantic contexts, and change points mark when these transformations occurred most dramatically.

4. Results and analysis

Our analysis employs the three analytical functions of DBE described in Section 3.3: absolute drift, embedding

neighborhoods, and change points. First, we calculated an ‘absolute drift,’ which reflects the difference in each word’s semantic representation between the earliest and the latest time slices. This step enables identifying words that underwent the most significant changes in meaning. However, absolute drift measures only the *magnitude* of change, not its direction.

Because absolute drift alone does not capture the nature of semantic shifts, we also analyzed embedding neighborhoods to map how a word’s closest semantic associations evolve over time—calculating the ‘embedding neighborhoods’ and the ‘change points’ of specific words. By examining how a word’s closest semantic associates changed across different time periods, we traced the shifting contexts surrounding high-drift words. In addition, we compared these computed outcomes to present the contrast and temporal progress of the semantic surroundings of the term ‘Soviet Union’ in both newspapers.

Finally, to further contextualize these shifts, we conducted close reading of articles containing high-drift keywords in selected time slices. This step offers a concrete yet revealing interpretation of the patterns observed in the previous steps. It shows how specific words were deployed in media discourse, which historical events they were tied to, and whether their shifts corresponded to broader geopolitical realignments.

4.1 Absolute drifts

In the RMRB top 100 list, the high drift of ‘苏联’ (Soviet Union; sixth) underscores the shifting portrayal of this particular superpower in Chinese media. The list contains other terms directly evocative of the Cold War’s ideological contours. Words such as ‘斗争’ (struggle; 15th), ‘社会主义’ (socialism; 53rd), ‘集团’ (bloc; 54th), and ‘革命’ (revolution; 62nd) with high drift values, point to a dynamic ideological discourse within RMRB. Furthermore, the presence of major powers like ‘美国’ (United States; 90th), ‘英国’ (United Kingdom; 48th), ‘法国’ (France; 70th), and ‘日本’ (Japan; 67th) among the top drift words shows the geopolitical dimension of the observed semantic shifts. These states, representing key actors in the post-1945 international order, suggest that RMRB’s portrayal of their roles and relationships with China was not static.

A similar Cold War influence is evident in the top 100 drift list for CSIB, though with a distinct focus. ‘북한’ (North Korea; 14th) emerges as a central keyword, underscoring the persistent yet dynamic relevance of North Korea in South Korean media discourse. Other high-drift terms such as ‘중국’ (China; 35th), ‘중공’ (Communist China; 56th), ‘독일’ (Germany; 84th), and

‘유엔’ (United Nations; 94th) reflect South Korea’s entanglement in the global Cold War framework. Interestingly, ‘소련’ (189th) does not appear in the top 100, but its drift within the larger dataset (2,432 words analyzed) suggests that Soviet-related discourse in CSIB also underwent considerable shifts, albeit to a lesser extent than in RMRB.

A direct comparison of the RMRB and CSIB lists reveals that the most salient divergence lies in the central referents of their semantic shifts. RMRB’s list is conspicuously centered around the ‘苏联’ and a constellation of ideology-related terms. In contrast, CSIB’s list, while also not irrelevant to the ongoing Cold War, is seemingly dominated by peninsula-centric concerns. ‘북한’ stands out as a high-drift keyword, overshadowing any single external power. The presence of ‘원조’ (aid; 37th) and ‘지원’ (support; 97th) suggests that the context of foreign assistance was arguably redefined as South Korea’s relationship with the US changed over time. The Cold War’s two main axes—‘미국’ (United States; 346th) and ‘소련’ (189th)—are less prominent in CSIB’s data, suggesting a different emphasis in South Korean discourse.

It is worth noting that these comparisons provide only a preliminary view of the larger picture. Absolute drift only measures the *magnitude* of change. While it would be valuable to examine the full semantic neighborhoods of all high-drift words, constraints on space and computational resources necessitate narrowing down a target term. We therefore focused on the Soviet Union, because it serves as a revealing case for comparing the semantic shifts in both PRC and ROK media discourse. With the Soviet Union as a common reference point in both corpora, we can computationally compare how media discourse in two ideologically opposed Cold War states shaped the national opinion while adapting to changing geopolitical realities.

4.2 Embedding neighborhoods

To further explore ideological and geopolitical shifts in Cold War discourse, we analyzed the embedding neighborhoods of the ‘Soviet Union’ within four key time slices: 1951–5, 1961–5, 1966–70, and 1986–90 (Table 3). The first slice, 1951–55, corresponds to the Korean War (1950–3) and its aftermath, a defining event that entrenched conflicts in East Asia and beyond. The second slice, 1961–5, marks the Sino-Soviet split, which had profound consequences for China’s ideological positioning. The Cuban Missile Crisis (1962) exacerbated tensions, as Beijing criticized Moscow’s perceived capitulation to US pressure. Meanwhile, in South Korea, the 1961 coup led by Park Chung-hee initiated a shift toward authoritarian rule, with heavy state control over

the media. The third slice, 1966–70, coincides with the Cultural Revolution (1966–76) in China, which sought to eradicate ‘revisionist’ influences, stemming from the Soviet Union, and led to an escalation of revolutionary rhetoric (Gilkison and Kurzynski 2024). For both the PRC and the ROK, the Vietnam War, or the Second Indochina War, was partly an extended national conflict. The final slice, 1986–90, marks a period of political transformation and ideological recalibration in both China and South Korea. In Beijing, Deng Xiaoping’s economic reforms had taken full effect, shifting the dominant discourse from revolutionary struggle to pragmatic development. In Seoul, this period saw the June Democracy Movement (1987), which led to the end of military dictatorship and the transition to a democratic government that pursued the Nordpolitik, an initiative of appeasement toward the broader socialist world.

Table 3 traces the shifting embedding neighborhoods of the ‘Soviet Union’ in RMRB and CSIB. While each newspaper likely framed its own nation deliberately, there was hardly any direct coordination between the two dailies in framing the Soviet Union. Comparing these semantic constellations thus offers insights into how mainstream media in each societal context discursively positioned the Soviet Union. Predictable associations, such as ‘Russia,’ ‘Moscow,’ and the ‘Soviet government,’ emerge, but the other surrounding terms warrant closer scrutiny.

In the 1951–5 slice, both RMRB and CSIB associate the ‘Soviet Union’ with diplomatic language. However, CSIB also links it to terms, such as ‘물질’ (material) and ‘폭력’ (violence), indicating that, in the South Korean context, the Soviet Union was semantically entangled with more negative or tangible concerns. RMRB’s neighborhood during this period centers on diplomatic engagement, with terms like ‘外交使节’ (diplomatic envoy), ‘临时代办’ (interim diplomatic agent), ‘参赞’ (diplomatic attaché), and ‘中国大使馆’ (Chinese Embassy), underscoring a focus on formal state relations. On the other hand, CSIB’s neighborhood is more geographically expansive. Terms, including ‘서방’ (the West, Western), ‘아프리카’ (Africa), ‘유럽’ (Europe), and ‘아시아’ (Asia), suggest a broader mapping of global influence. Although DWE identifies patterns, it does not fully account for them; a close reading of the source articles is required to interpret their meaning. This divergent framing of the Soviet Union will be discussed further in the next section.

These patterns reveal an interesting divergence in how RMRB and CSIB framed the Soviet Union during the early 1960s, a period defined by the intensifying Sino-Soviet split. RMRB emphasizes the Soviet Union’s global entanglements in ideological conflict, particularly across

the socialist bloc and postcolonial contexts. The prominence of ‘南斯拉夫’ (Yugoslavia)—ranked first—reflects its fraught relations with Moscow after their ideological rupture and unsuccessful rapprochements (Kramer 2017a,b). The inclusion of ‘捷克斯洛伐克’ (Czechoslovakia) anticipates the Prague Spring’s unrest later in the decade (Applebaum 2019). Meanwhile, terms like ‘葡萄牙’ (Portugal) and ‘科威特’ (Kuwait) point to China’s view of Soviet influence beyond Eastern Europe, referencing Soviet support for anti-colonial movements in Portuguese Africa (Telepneva 2022) and opposition to Kuwait’s UN membership (Prantner 2005). This framing suggests a Chinese concern with the Soviet Union’s contested role in both global socialism and anti-imperialist struggles. In contrast, CSIB adopts a more security-oriented perspective, primarily portraying the Soviet Union as a regional threat. The strong association with ‘북한’ reflects perceptions of Moscow as Pyongyang’s patron, reinforcing its image as a direct threat to South Korea. The inclusion of ‘중공’ further underscores the perceived strategic danger posed by the Sino-Soviet relations in the Korean Peninsula. In sum, RMRB situates the Soviet Union within larger diplomatic and ideological realignments, while CSIB foregrounds national security, emphasizing the Soviet role in sustaining adversarial regimes near South Korea’s borders.

From 1966 to 1970, embedding neighborhoods in RMRB and CSIB reflect a deepening divergence in portrayals of the Soviet Union, shaped by the radicalization of Chinese rhetoric during the early Cultural Revolution and South Korea’s continued security concerns. RMRB increasingly framed the USSR as a revisionist betrayer of socialist ideals, evidenced by emotionally charged terms like ‘奇怪’ (strange; baffled) and recurring references to ‘赫鲁晓夫’ (Nikita Khrushchev) despite his ouster, suggesting a sustained critique of his ideological legacy (Torigian 2022, Wagner 2023). The prominence of ‘南斯拉夫,’ ‘捷克斯洛伐克,’ and ‘扩张’ (expansion) underscores a narrative of ideological deviation and geopolitical rivalry within the socialist bloc. Meanwhile, CSIB emphasized the Soviet Union’s militaristic threat through terms like ‘ 침략’ (invasion), and ‘무력’ (military force), framing the USSR as a primary actor in the communist security threat to South Korea. References to ‘독일’ and ‘서방’ also signal CSIB’s perception of the Soviet Union as entangled in Cold War power dynamics, particularly around German unification. While both newspapers depict the Soviet Union as aggressive, RMRB casts it as an ideological traitor, whereas CSIB portrays it as a geopolitical and military adversary.

By the late 1980s, Deng Xiaoping’s economic reforms had fundamentally altered China’s priorities, shifting away from ideological confrontation. The presence of ‘波

兰’ (Poland) could be linked to the rise of Solidarity and national movements in Eastern Europe, signaling a shift in global socialist discourse. In the CSIB, the term ‘empire’ was often referred to in relation to the Eastern Bloc, a major Soviet sphere of influence (Zubok 2009).

4.3 Detecting change points

While absolute drift measures the semantic distance between the earliest and latest time slices, DWE allows for a more granular decomposition of this cumulative change through the detection of change points—that is, a period-by-period measurement of how much a word’s meaning shifts between consecutive intervals. These values unpack *when* and *how evenly* semantic change is distributed across time. Figure 3 presents the change point trajectories for ‘苏联’ in RMRB and ‘소련’ in CSIB. Each value represents the magnitude of semantic shift from the preceding five-year period, with higher values indicating sharper transformations in how the term was contextually deployed compared to the previous slice.

RMRB shows high volatility early on, peaking at 1.78 in 1956–60 and remaining elevated until 1986–90. Yet, the semantic change value drops sharply afterward, reaching a low of 0.93 in 1991–5. This drop indicates a rapid stabilization of the term’s meaning following the Soviet Union’s collapse. With the referent no longer present as a geopolitical actor, its semantic dynamism in the corpus diminished.

In contrast, CSIB’s change values remained relatively consistent throughout (between 1.08 and 1.29). This trend does not imply semantic stagnation. Rather, the meaning of ‘소련’ continued to shift steadily even after the Cold War’s end, likely reflecting South Korea’s security concerns regarding Russia and North Korea, where the Soviet Union’s successor state and its satellite remained immediate threats.

The methodological value of change point detection lies in its capacity to channel interpretive attention into specific temporal windows where semantic transformation was most pronounced. In this sense, change point detection serves as a heuristic for identifying historical intervals where meanings shifted more or less intensely, thereby bridging large-scale computational results with historically grounded, close-reading analysis.

4.4 Close reading

In examining the embedding neighborhoods, we turn our attention to keywords whose semantic shifts do not lend themselves to intuitive explanations. While the term ‘Soviet Union’ moved from diplomatic to adversarial contexts in RMRB coverage, the presence of seemingly

We emphasize the complementary strengths of DWE; whereas it identifies shifts in meaning, close reading is necessary to clarify the causes and developments of each instance. DWE raises more questions than it answers, and therein lies its potential as a productive research tool. DWE offers departure points for research (Drucker 2017), guiding scholars to raise more meaningful questions regarding the correlations, patterns, or forces embedded within. By combining the *scale* of ‘distant reading’ with the *granular precision* of close reading, we claim that our approach demonstrates how to expand the scope of inquiry while maintaining interpretive rigor.

Another strength of this methodology lies in its potential for cross-lingual and longitudinal comparisons. As more digitized newspaper archives become available, particularly in underrepresented languages (LeBlanc 2024), researchers can analyze at an unprecedented scale how different societies framed the actors or events of geopolitical importance over time. If multiple corpora cover the same period, DWE enables comparisons of how key terms evolved across varied lingual contexts. This capability is particularly relevant given the increasing interest in digitized newspapers as a ‘new Eldorado’ for historians (Bunout et al., 2022). Yet, as with all large-scale digital approaches, scholars must navigate this vast corpus with care. By illustrating the use of DWE in cross-linguistic research, our study provides a promising model for digital humanities researchers to engage with the emerging ‘Eldorado’ without getting lost in computational analysis.

6. Conclusion

Our study has several limitations. First, it depends on the availability of digitized archives. Our focus on RMRB and CSIB was possible only because extensive, longitudinal datasets for these newspapers were available, highlighting the critical role of digitization in print media research. Without such archives, similar studies may face significant constraints.

Second, the findings are sensitive to methodological choices in preprocessing and modeling. Decisions about text cleaning, segmentation, and the treatment of linguistic nuances, such as synonyms or transliterations, can reproduce underlying data biases, influencing word embeddings and their interpretation. We therefore present our method as a heuristic framework open to further refinement.

Third, the scope of our corpus is limited to two major newspapers, both of which reflect dominant perspectives. While these sources offer insights into official

narratives, they omit alternative or dissenting voices. A fuller account of Cold War discourse in East Asia would require expanding the corpus to include other newspapers and media from a broader spectrum of ideological and regional contexts.

Finally, we acknowledge that Cold War media discourse was not purely textual. Visual media, such as photographs and political cartoons, played a substantial role in shaping public perception (McNair 1988). Recent works in computer vision demonstrate the feasibility of analyzing image data from newspapers and magazines (Fyfe and Ge 2018; Du et al., 2023; Soh et al., 2023). The integration of multimodal analysis would thus yield a more well-rounded understanding of how discourse was produced during the Cold War.

Acknowledging these constraints, we propose several directions for future research. Expanding the corpus to include additional contemporary newspapers, such as other Chinese and Korean publications, as well as Soviet and Western outlets, would broaden the comparative scope and incorporate perspectives beyond official narratives. Integrating other computational techniques, such as moral attitude analysis (Chen et al., 2024), could also help quantify shifts in tone and framing. Moreover, the method demonstrated here is adaptable to other historical periods. Applying similar analyses to media coverage during major global events, such as the World Wars (Martin 2022; Buchanan 2023), could test the method’s generalizability and shed light on how public narratives shift across discursive contexts.

Despite technical challenges, we present a method that extracts insights while minimizing distortion in newspaper data. By accounting for the linguistic particularities of Chinese and Korean, our approach contributes to ongoing discussions on maximizing the value of data in digital scholarship in the humanities and social sciences. We hope that these findings will encourage further research into the evolution of words and concepts within specific linguistic environments, helping to bridge the gap between computational methods and humanities research.

Author contributions

Byungjun Kim (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review & editing), Jonghyun Jee (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original

draft, Writing—review & editing), and Donghyun Woo (Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Writing—review & editing)

Conflicts of interest. The authors declare no competing financial or non-financial interests.

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2023-00241749), and also by Woojung Historical System Project funded by the Korea Advanced Institute of Science and Technology Development Fund (B0101001553).

Data availability

The trained data, code, and the instructions for replicating the results of this study are available at the link: <https://figshare.com/s/9173a4c409962868e45b>.

Notes

1. The entire data preprocessing and modeling process was conducted using Python, and the corresponding code is publicly available on the Figshare website.
2. <https://newslibrary.chosun.com>
3. <https://github.com/foxsjy/jieba>
4. <https://radimrehurek.com/gensim>
5. We developed a module based on the existing package (https://github.com/llefebure/dynamic_bernoulli_embeddings) to train large-scale data using PyTorch's Distributed Data Parallel (DDP). By enabling the use of multiple GPUs, this module allows for efficient training, making it possible for other researchers to train models more effectively. The developed module is available at: https://github.com/ByungjunKim/dynamic_bernoulli_embeddings
6. See RMRB, '多尔蒂科斯、恩克鲁玛和塞拉西一世在不结盟国家会议上讲话 强烈谴责新老殖民主义和帝国主义,' 4 Sep. 1961. <https://www.laoziliao.net/rmrb/1961-09-04-3#276374>, accessed 5 Mar. 2025.
7. See CSIB, '危機克服은實力에서,' 1 Dec. 1953. https://newslibrary.chosun.com/view/article_view.html?id=920019530111m10127&set_date=19530111&page_no=1, accessed 5 Mar. 2025.

References

- Applebaum, R.** (2019) *Empire of Friends: Soviet Power and Socialist Internationalism in Cold War Czechoslovakia*. Ithaca, NY: Cornell University Press.
- Bamler, R., and Mandt, S.** (2017) 'Dynamic word embeddings', in *Proceedings of the 34th International Conference on Machine Learning (ICML'17)*, pp. 380–89. Sydney, Australia: PMLR. <https://proceedings.mlr.press/v70/bamler17a.html>, accessed 13 Oct. 2024.
- Bloomfield, L.** (1933) *Language*. New York, NY: Henry Holt and Company.
- Buchanan, A.** (2023) 'Globalizing the Second World War', *Past and Present*, 258: 246–81. <https://doi.org/10.1093/pastj/gtab042>
- Bunout, E., Ehrmann M., and Clavert F., eds** (2022) *Digitised Newspapers—A New Eldorado for Historians?: Reflections on Tools, Methods and Epistemology*. Berlin: De Gruyter.
- Card, D. et al.** (2022) 'Computational Analysis of 140 Years of US Political Speeches Reveals More Positive but Increasingly Polarized Framing of Immigration', *Proceedings of the National Academy of Sciences*, 119: 1–9. <https://doi.org/10.1073/pnas.2120510119>
- Chen, A. X., Sun, S., and Yu, H.** (2024) 'Moral Attitudes towards Effort and Efficiency: A Comparison between American and Chinese History', *Humanities and Social Sciences Communications*, 11: 1–14. <https://doi.org/10.1057/s41599-024-03603-3>
- Costa, L. F.** (2018) 'A Method for Content Analysis Applied to Newspaper Coverage of Japanese Personalities in Brazil and Portugal', *Digital Scholarship in the Humanities*, 33: 231–247. <https://doi.org/10.1093/llc/fqx050>
- Das Gupta, A., and Lüthi L., eds** (2017) *The Sino-Indian War of 1962: New Perspectives*. London: Routledge.
- Di Gennaro, G., Buonanno, A., and Palmieri, F. A. N.** (2021) 'Considerations about Learning Word2Vec', *The Journal of Supercomputing*, 77: 12320–35. <https://doi.org/10.1007/s11227-021-03743-2>
- Drucker, J.** (2017) 'Why Distant Reading Isn't', *PMLA/Publications of the Modern Language Association of America*, 132: 628–35. <https://doi.org/10.1632/pmla.2017.132.3.628>
- Du, L., Le, B., and Honig, E.** (2023) 'Probing Historical Image Contexts: Enhancing Visual Archive Retrieval through Computer Vision', *Journal on Computing and Cultural Heritage*, 16: 1–17. <https://doi.org/10.1145/3631129>
- Fainberg, D.** (2021) *Cold War Correspondents: Soviet and American Reporters on the Ideological Frontlines*. Baltimore, MD: Johns Hopkins University Press.

- Firth, J. R.** (1957) *Papers in Linguistics, 1934-1951*. London: Oxford University Press.
- Friedman, J. S.** (2015) *Shadow Cold War: The Sino-Soviet Competition for the Third World*. Chapel Hill, NC: University of North Carolina Press.
- Fyfe, P., and Ge, Q.** (2018) 'Image Analytics and the Nineteenth-Century Illustrated Newspaper', *Journal of Cultural Analytics*, 3: 1–25. <https://doi.org/10.22148/16.026>
- Garg, N. et al.** (2018) 'Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes', *Proceedings of the National Academy of Sciences*, 115: 3635–44. <https://doi.org/10.1073/pnas.1720347115>
- Ghasiya, P., and Okamura, K.** (2021) 'Understanding the Middle East Through the Eyes of Japan's Newspapers: A Topic Modelling and Sentiment Analysis Approach', *Digital Scholarship in the Humanities*, 36: 871–85. <https://doi.org/10.1093/llc/fqab019>
- Gilkison, A., and Kurzynski, M.** (2024) 'Vectors of Violence: Legitimation and Distribution of State Power in the People's Liberation Army Daily (Jiefangjun Bao)', *Journal of Cultural Analytics*, 9, 1956–89. <https://doi.org/10.22148/001c.115481>
- Guldi, J.** (2024) 'The Revolution in Text Mining for Historical Analysis is Here', *The American Historical Review*, 129: 519–43. <https://doi.org/10.1093/ahr/rhae163>
- Gulordava, K., and Baroni, M.** (2011) 'A Distributional Similarity Approach to the Detection of Semantic Change in the Google Books Ngram Corpus', in Pado S., and Peirsman Y. (eds) *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics (GEMS'11)*, pp. 67–71. Edinburgh, UK: Association for Computational Linguistics. <https://aclanthology.org/W11-2508>, accessed 14 Sept. 2024.
- Hamilton, W. L., Leskovec, J., and Jurafsky, D.** (2016) 'Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change', in Su J., Duh K., and Carreras X. (eds) *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP'16)*, pp. 2116–21. Austin, Texas: Association for Computational Linguistics. <http://aclweb.org/anthology/D16-1229>, accessed 15 Sept. 2024.
- Han, J., and Lee, G.** (2013) 'A Comparative Study of the Accuracy of Quotation-Embedded Headlines in Chosun Ilbo and the New York Times from 1989 to 2009', *Korea Journal*, 53: 65–90.
- Hengchen, S. et al.** (2021) 'A Data-Driven Approach to Studying Changing Vocabularies in Historical Newspaper Collections', *Digital Scholarship in the Humanities*, 36: ii109–26. <https://doi.org/10.1093/llc/fqab032>
- Hilpert, M.** (2008) *Germanic Future Constructions: A Usage-Based Approach to Language Change*. Amsterdam: John Benjamins Publishing Company.
- Jersild, A.** (2014) *The Sino-Soviet Alliance: An International History*. Chapel Hill, NC: The University of North Carolina Press.
- Jiang, Z.** (2024) 'Mining Themes, Emotions, and Stance in the News Coverage of the Russia–Ukraine War from Reuters and Xinhua', *Digital Scholarship in the Humanities*, 39: 609–24. <https://doi.org/10.1093/llc/fqae015>
- Jurafsky, D., and Martin, J. H.** (2000) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Hoboken, NJ: Prentice Hall.
- Kim, H. S.** (2018) 'Korean Journalism: From Partners of Political Power to Adversarial Agents of Social Change', in D. Y. Jin and N. J. Kwak (eds) *Communication, Digital Media, and Popular Culture in Korea: Contemporary Research and Future Prospects*, pp. 125–52. Lanham, MD: Lexington Books.
- Kirilloff, G.** (2022) 'Computation as Context: New Approaches to the Close/Distant Reading Debate', *College Literature*, 49: 1–25.
- Kramer, M.** (2017a) 'The Changing Pattern of Soviet–East European Relations 1953–1968', in N. Naimark, S. Pons, and S. Quinn-Judge (eds) *The Cambridge History of Communism: Volume 2, The Socialist Camp and World Power 1941–1960s*, pp. 139–69. Cambridge: Cambridge University Press.
- Kramer, M.**, (2017b) 'Stalin, the Split With Yugoslavia, and Soviet–East European Efforts to Reassert Control, 1948–53', In: S. Rajak et al (eds) *The Balkans in the Cold War*, pp. 29–63. New York, NY: Springer.
- Kutuzov, A. et al.** (2018) 'Diachronic Word Embeddings and Semantic Shifts: A Survey', in Bender E. M., Derczynski L., and Isabelle P. (eds) *Proceedings of the 27th International Conference on Computational Linguistics (COLING'18)*, pp. 1384–97. Santa Fe, New Mexico: Association for Computational Linguistics <https://aclanthology.org/C18-1117/>, accessed 23 Oct. 2024.
- LeBlanc, Z.** (2024) 'More Than Keywords: Histories of Decolonization and Digitized Newspapers', *The American Historical Review*, 129: 164–168. <https://doi.org/10.1093/ahr/rhad499>
- Lee, M.** (2024) 'Kiwi: Developing a Korean Morphological Analyzer Based on Statistical Language Models and Skip-Bigram', *Korean Journal of Digital Humanities*, 1: 109–36.
- Lenoe, M. E.** (2004) *Closer to the Masses: Stalinist Culture, Social Revolution, and Soviet Newspapers*. Cambridge, MA: Harvard University Press.
- Li, H.** (2024) *Fighting on the Cultural Front: U.S.-China Relations in the Cold War*. New York, NY: Columbia University Press.

- Martin, J.** (2022) 'Globalizing the History of the First World War: Economic Approaches', *The Historical Journal*, 65: 838–85. <https://doi.org/10.1017/s0018246x21000765>
- Masuda, H.** (2015) *Cold War Crucible: The Korean Conflict and the Postwar World*. Cambridge, MA: Harvard University Press.
- McNair, B.** (1988) *Images of the Enemy: Reporting the New Cold War*. New York, NY: Routledge.
- Michel, J.-B., Google Books Team. et al.** (2011) 'Quantitative Analysis of Culture Using Millions of Digitized Books', *Science (New York, N.Y.)*, 331: 176–182. <https://doi.org/10.1126/science.1199644>
- Mikolov, T. et al.** (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv : 1301.3781, preprint: not peer reviewed.
- Moretti, F.** (2013) *Distant Reading*. London: Verso Books.
- Nicholson, B.** (2013) 'The DIGITAL TURN: Exploring the Methodological Possibilities of Digital Newspaper Archives', *Media History*, 19: 59–73. <https://doi.org/10.1080/13688804.2012.752963>
- Pang, N., Sun, M., and Zhu, H.** (2024) 'Representation of Migrant Workers in Chinese Newspapers: combining Topic Modeling and Critical Discourse Analysis', *Digital Scholarship in the Humanities*, 39: 1092–111. <https://doi.org/10.1093/llc/fqae062>
- Patel, K., and Bhattacharyya, P.** (2017) 'Towards Lower Bounds on Number of Dimensions for Word Embeddings', in Kondrak G., and Watanabe T. (eds) *Proceedings of the Eighth International Joint Conference on Natural Language Processing (IJCNLP'17)*, pp. 31–36. Taipei, Taiwan: Asian Federation of Natural Language Processing. <https://aclanthology.org/I17-2006/>, accessed 1 July 2025.
- Pennington, J., Socher, R., and Manning, C.** (2014) 'GloVe: Global Vectors for Word Representation', in Moschitti A., Pang B., and Daelemans W. (eds) *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP'14)*, 1532–43. Doha, Qatar: Association for Computational Linguistics. <http://aclweb.org/anthology/D14-1162>, accessed 12 Sept. 2024.
- Prantner, Z.** (2005) 'Socialist Standpoint in the Kuwait Question, 1961-1963', *Études Sur La Région Méditerranéenne*, 14: 51–66.
- Radchenko, S.** (2009) *Two Suns in the Heavens: The Sino-Soviet Struggle for Supremacy, 1962-1967*. Stanford, CA: Stanford University Press.
- Rudolph, M., and Blei, D.** (2018) 'Dynamic embeddings for language evolution', in *WWW 2018: The 2018 Web Conference*, pp. 1003–11. Lyon, France: ACM Press. <https://www.cs.columbia.edu/~blei/papers/RudolphBlei2018.pdf>, accessed 30 Oct. 2025.
- Sagi, E., Kaufmann, S., and Clark, B.** (2009) 'Semantic Density Analysis: Comparing Word Meaning across Time and Phonetic Space', in Basili R., and Pennacchiotti M. (eds) *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics (GEMS'09)*, pp. 104–11. Athens, Greece: Association for Computational Linguistics. <https://aclanthology.org/W09-0214>, accessed 14 Sept. 2024.
- Soh, L. et al.** (2023) 'Applying Image Analysis and Machine Learning to Historical Newspaper Collections', *The American Historical Review*, 128: 1382–1389. <https://doi.org/10.1093/ahr/rhad369>
- Tahmasebi, N. et al.** (2021) *Computational Approaches to Semantic Change*. Berlin: Language Science Press.
- Telepneva, N.** (2022) *Cold War Liberation: The Soviet Union and the Collapse of the Portuguese Empire in Africa, 1961–1975*. Chapel Hill, NC: The University of North Carolina Press.
- Tinn, H.** (2025) *Island Tinkerers: Innovation and Transformation in the Making of Taiwan's Computing Industry*. Cambridge, MA: The MIT Press.
- Torigian, J.** (2022) "'You Don't Know Khrushchev Well": The Ouster of the Soviet Leader as a Challenge to Recent Scholarship on Authoritarian Politics', *Journal of Cold War Studies*, 24: 78–115. https://doi.org/10.1162/jcws_a_01043
- Traugott, E. C.** (2017) 'Semantic Change', in: M. Aronoff (ed) *Oxford Research Encyclopedia of Linguistics*, pp. 385–406. New York, NY: Oxford University Press.
- Wagner, M.** (2023) 'Excoriating Stalin, Criticizing Mao: Entangled Reevaluations of the Past in the 1950s Soviet Union and 1970s/80s China', *The American Historical Review*, 128: 1105–1143. <https://doi.org/10.1093/ahr/rhad238>
- Westad, O. A.** (2005) *The Global Cold War: Third World Interventions and the Making of Our Times*. New York: Cambridge University Press.
- Wevers, M.** (2019) 'Using Word Embeddings to Examine Gender Bias in Dutch Newspapers, 1950-1990', in Tahmasebi N. et al. (eds) *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change (LChange'19)*, pp. 92–97. Florence, Italy: Association for Computational Linguistics. <https://www.aclweb.org/anthology/W19-4712>, accessed 13 Oct. 2024.
- Wijaya, D. T., and Yeniterzi, R.** (2011) 'Understanding Semantic Change of Words over Centuries', in *Proceedings of the 2011 International Workshop on DETecting and Exploiting Cultural diversity on the Social Web (CIKM '11)*, pp. 35–40. Glasgow, UK: Association for Computing Machinery. <https://dl.acm.org/doi/10.1145/2064448.2064475>, accessed 14 Sept. 2024.

- Wijffes, H.** (2017) 'Digital Humanities and Media History: A Challenge for Historical Newspaper Research', *TMG Journal for Media History*, 20: 4–24. <https://doi.org/10.18146/2213-7653.2017.277>
- Yi, G.** (2013) 'The "Propaganda State" and Sino-American Rapprochement: Preparing the Chinese Public for Nixon's Visit', *The Journal of American-East Asian Relations*, 20: 5–28. <https://doi.org/10.1163/18765610-02001005>
- Zubok, V. M.** (2009) *A Failed Empire: The Soviet Union in the Cold War from Stalin to Gorbachev*. Chapel Hill, NC: The University of North Carolina Press.