The role of English Wikipedia in mediating East Asian historical disputes: the case of Balhae

Jonghyun Jee, Byungjun Kim & Bong Gwan Jun

To cite this article: Jonghyun Jee, Byungjun Kim & Bong Gwan Jun (22 Apr 2024): The role of English Wikipedia in mediating East Asian historical disputes: the case of Balhae, Asian Journal of Communication, DOI: 10.1080/01292986.2024.2342822

To link to this article: https://doi.org/10.1080/01292986.2024.2342822

Published online: 22 Apr 2024.

Submit your article to this journal

View related articles

View Crossmark data
ABSTRACT

With the expansion of the Internet in the early twenty-first century, public historical disputes between South Korea and China have become more pronounced. While existing research has focused on domestic Korean or Chinese digital spaces, the convergence or clash of these national narratives on global platforms remains underexplored. Our research bridges this gap by investigating the role of online platforms in shaping public historical debates. We compare the portrayals of Balhae, an ancient kingdom with contested contexts between the two nations. By comparing Chinese, Korean, and English Wikipedia entries on Balhae, we identify differences in narrative construction and framing. Employing Habermas’s typology of human action, we scrutinize related talk pages on English Wikipedia to examine the strategic actions multinational contributors employ to shape historical representation. This exploration reveals the dual role of online platforms in both amplifying and mediating historical disputes. While Wikipedia’s policies promote rational discourse, our findings indicate that contributors often vacillate between strategic and communicative actions. Nonetheless, the resulting article approximates Habermasian ideals of communicative rationality. This interaction accentuates the challenges and potential of fostering rational online discourse, the complexities of digital collaboration, and the quest for neutrality in open, decentralized platforms.

ARTICLE HISTORY

Received 10 August 2023
Revised 19 March 2024
Accepted 4 April 2024

KEYWORDS

Wikipedia; digital nationalism; neutrality; communicative action; strategic action; Balhae; Sino-Korean relations

Introduction

In East Asia, the discussion of history has found new frontiers in online platforms. This transformation is evident in the contested historical narratives emerging in the domestic cyberspace of South Korea and China (Chase, 2011; Chung, 2008; Hundt & He, 2015; Ke, 2022; Lee & Yu, 2015; Schneider, 2018). The emergence of Web 2.0, characterized by user-generated content and platforms such as Wikipedia, has reshaped the terrain of historical engagement and introduced unexplored complexities (Manfra et al., 2009). Beyond the confines of traditional textbooks or state-controlled media
(Schäfer, 2015), these decentralized platforms foster a dynamic, interactive, and often contentious discourse on history. This digital evolution has opened doors to a rich yet intricate dialogue around historical representation, highlighting both challenges and opportunities in mediating historical disputes (Fuchs, 2019; Schneider, 2022). While domestic Korean and Chinese cyberspace has been extensively researched, a discernible gap remains in exploring how these national narratives intersect and conflict on shared online platforms.

The intertwining of divergent narratives in online spaces such as Wikipedia raises questions about neutrality, collective memory, and implications for multinational, multilingual conflicts. Previous research has probed various controversies – including the Srebrenica massacre (Rogers & Sendijarevic, 2012), the Ukrainian Insurgent Army and the Latvian Legion (Kaprans, 2016), the Battle of Lviv (Makhortykh, 2017), Egyptian Revolution (Ferron & Massa, 2011) and others – but the context of Wikipedia in East Asia remains relatively unexplored. This paper contributes to the burgeoning field by focusing on fostering rational discourse amidst growing historical contention in the East Asian context. It situates the study within the broader discourse of historical disputes, such as the Istanbul/Ístanbul controversy (Jones, 2018), and the Falklands/Malvinas War (Oeberst et al., 2020).

This study has two objectives:

1. To compare the entry for Balhae1 – an ancient kingdom with contested historical interpretations (Crossley, 2016; Kim, 2016; Reckel, 2015; Sloane, 2014; Song, 1990) – in the Chinese, Korean, and English versions of Wikipedia, identifying differences in narrative construction and framing.

2. To investigate related talk pages in English Wikipedia, applying the Habermasian typology of human action to reveal how multinational contributors negotiate historical representation.

Guided by Jürgen Habermas’s *Theory of Communicative Action* (1984, 1987), this research employs a typology of human actions to interpret the underlying processes and motivations of Wikipedia contributors. The three primary types of action – instrumental, strategic, and communicative – are mapped onto Wikipedia interactions, establishing a conceptual lens that aligns with existing literature applying the Habermasian model to Wikipedia (Hansen et al., 2009; see also Barton, 2005; Bubendorff et al., 2021; Firer-Blaess, 2011; Fuchs, 2015; Goldspink, 2010; Kopf, 2022). Critical questions arise from this alignment, such as whether Wikipedia contributors embrace communicative action in pursuing neutrality, or if strategic actions covertly further personal viewpoints. On top of the qualitative analysis, we also employ ‘XTools,’ an external data collection tool, to track quantitative information about the revision history and contributors of the articles, aiding in understanding the collaborative process and identifying the primary contributors.

This study underscores the possibilities and challenges of promoting rational discourse in decentralized online environments. By analyzing the interactions among multinational users in mediating historical disputes, our findings contribute to a deeper understanding of how neutrality is attained in collaborative knowledge production platforms (Bryant et al., 2005). Furthermore, this study might inform future efforts to
encourage more inclusive discussions, not just on Wikipedia itself, but anywhere people come together to build a shared understandings of the past.

**Theoretical framework**

We aim to investigate the potential of online platforms, with a specific focus on Wikipedia, to cultivate constructive and inclusive discussions. We term these discussions ‘rational discourse,’ which, for the purposes of this study, refers to dialogues that are reasoned, objective, and aimed at mutual understanding. Our theoretical foundation is rooted in Habermas’s *Theory of Communicative Action* (1984, 1987), which posits that societal progress hinges on the ability to engage in such rational discourse. Some researchers, including Hansen et al. (2009), contend that Wikipedia closely aligns with the prerequisites for fostering rational discourse. This perspective prompts us to question: Can the English Wikipedia truly function as a neutral arena for deliberating contentious historical narratives?

To explore the dynamics of these interactions, we turn to Habermasian typology of human actions. This typology offers insights into the processes and motivations that drive human social interactions. Habermas delineates three primary types of action: instrumental, strategic, and communicative.

**Instrumental action** refers to goal-oriented activities wherein an actor uses specific means to achieve a desired end (Habermas, 1984, p. 285). The action is determined by the calculation of the most efficient way to achieve the objective, and does not involve other actors or their intentions. In the context of Wikipedia, unilateral and undiscussed edits can be characterized as instrumental. The open-editing policy of Wikipedia invites such action – particularly from novice contributors who make edits without any social interaction or consideration of community guidelines.

**Strategic action**, like instrumental action, is also goal-oriented. However, it factors in other actors and how they are likely to respond. It is a type of social interaction in which actors try to get what they want by influencing others’ behavior (Habermas, 1984, p. 285). Imagine, for example, a new user whose edit gets reverted. The user soon discovers the importance of discussion on the ‘talk’ page. The user’s subsequent attempts to justify one’s edits are less a pursuit of neutrality than a propagation of one’s personal perspective. That user might invoke some policies to clinch an argument, or, in a more subtle manner, involves a level of maneuvering to achieve one’s goal. This type of behavior is strategic action, aimed at persuasion rather than collaboration.

**Communicative action**, conversely, refers to social interactions that strive for mutual understanding and consensus (Habermas, 1984, pp. 285–286.) Instead of focusing on individual goals, it emphasizes the principles of cooperation, consensus, and shared understanding through dialogue (Jemielniak, 2014). Thus, when contributors on platforms such as Wikipedia prioritize the development of a neutral encyclopedia and demonstrate a commitment to collaborative processes, their interactions can be deemed communicative action.

In examining user behaviors on Wikipedia, it becomes evident that Habermas’s typology offers a fitting lens. In particular, the motivations that govern Wikipedians engaged in discussions pose intriguing questions – Are they genuinely engaging in communicative action with neutrality as the shared goal? Or are their interactions strategic attempts...
to subtly promote personal viewpoints? We will explore these dynamics of strategic and communicative actions within Wikipedia’s talk pages.

**Neutrality in the context of Wikipedia**

The Neutral Point of View (NPOV) is a foundational principle at the heart of Wikipedia’s content creation and curation. This principle is pivotal for Wikipedia, ensuring that its vast repository of knowledge remains unbiased and representative. Wikipedia states that each article should represent ‘all the significant views that have been published by reliable sources on a topic’ presented in a fair and proportionate manner, thereby minimizing editorial bias.

The pursuit of neutrality, however, is not always straightforward. The NPOV frequently becomes a major flashpoint for contention in talk pages (Bilić, 2015, p. 1264; Rosenzweig, 2006, p. 122). The wording of the NPOV policy leaves room for interpretation. Questions arise: How does one determine which sources are reliable? Or decide which viewpoints deserve inclusion? Paradoxically, this inherent ambiguity in defining neutrality tends to mitigate conflicts among contributors (Matei & Dobrescu, 2011, p. 49). This also dovetails with Reagle’s observation that Wikipedians are more inclined to commonly accepted knowledge about a topic rather than engage in persuasion battles over correctness (2010, p. 53).

As with many ideals, the commitment to neutrality presents its challenges (Phillips, 2016, pp. 533–535). Some authors perceive their articles as adhering to NPOV, while others might contest this viewpoint. It is essential to recognize that these perceived discordances often arise not from deliberate bias but from unintentional misinterpretations or passionate disagreements (Reagle, 2010, p. 11). As Shirky (2009) astutely notes, Wikipedia’s content does not emerge from a harmonious collective hive mind. Instead, it is the product of ‘constant scrutiny and emendation,’ reflecting its inherently contentious and deliberative nature.

Within Wikipedia, a discernible tension exists between the NPOV and individual points of view (POV). This tension echoes Habermas’s distinction between communicative and strategic action. Hansen et al. (2009, p. 50) posit that while participants may enter discussions with subjective perspectives, Wikipedia’s structure nudges them to frame their arguments in line with NPOV. Thus, even if many contributors do not approach entries with purely communicative intentions (Barton, 2005), the resulting discourse leans toward rational communication by virtue of its guiding framework. Our interest lies in this facet of NPOV, where strategic actions to promote one’s POV can subtly masquerade as communicative actions.

**Local Wikipedia, global Wikipedia**

As of 2023, Wikipedia features a total of 334 language editions. The nature of these editions often mirrors the homogeneity of their user base: some predominantly showcase local narratives, while others gravitate towards a more global purview. Beyond mere linguistic variations, these editions differ in content emphasis, nuances, (Leshnick, 2022; Liao, 2009), and rule-making practices (Hwang & Shaw, 2022). Whereas most language editions correlate with a specific nation, languages such as English, Spanish, French,
Russian, and Arabic cater to broader global audiences. Notably, the English Wikipedia emerges as a nexus where diverse users converge to negotiate disparate perspectives (Massa & Scrinzi, 2012). Several works highlight how English Wikipedia is increasingly becoming a global knowledge hub, in contrast to smaller language Wikipedias that remain anchored in local values (Callahan & Herring, 2011; Hara et al., 2010).

And yet, a paradox persists within the English Wikipedia. While it stands as a global knowledge hub, it also struggles with the shadow of ‘information imperialism.’ A significant portion of English Wikipedia’s contributors hail from Western regions, predominantly North America and Europe. This demographic skew threatens to imbue the platform with a Western-centric bias, potentially marginalizing non-Western narratives. Avieson (2022) labels this phenomenon as information imperialism, highlighting instances in English Wikipedia where Western contributors mischaracterize Bhutanese cultures and overshadow indigenous perspectives.

Such Western bias is not always a deliberate act of cultural hegemony. Instead, it reflects the broader structural imbalances pervasive in the global context. Kumar’s research (2017) questions these inherent disparities in global knowledge dissemination. A case in point is the entry for the ‘Ganges’ river in the English Wikipedia. While the majority of Indians refer to the river as ‘Ganga,’ the title persists as ‘Ganges’ due to its more common use in standard English, namely British and American. This inclination towards Western narratives on Wikipedia indicates larger geopolitical undercurrents that include colonial histories, the global prominence of the English language, and uneven digital accessibility.

An intriguing illustration of this confluence is the development of the Balhae entry on English Wikipedia. In contrast to the two cases presented above, the primary contributors to the Balhae entry are not Westerners, but those who identify as South Korean, Chinese, Russian, or those of Korean or Chinese descent (see Table 2). Given this backdrop, our exploration focuses more on the possibilities and challenges of post-national forum than on the dimension of information imperialism.

**How each language edition frames ‘Balhae’**

Each Wikipedia language edition possesses some degree of autonomy over its rules, giving language communities the power to privilege certain sources and determine their cultural interpretations of notability and verification (Avieson, 2022, pp. 410–411). And yet, combined with this agency, the prevailing uniformity among users often leads to a monolithic, static portrayal of history. It is not unexpected that different narratives about the same topic appear in each language edition (e.g. the Falklands War in English, Guerra de las Malvinas in Spanish Wikipedia).

In a similar vein, to explore variations in narrative construction and framing, we examine the ‘Balhae’ entry across three Wikipedia language editions: Chinese, Korean, and English. While each associated country has its own online encyclopedias – such as Baidu Baike in China and Namuwiki in South Korea – we anchor our attention on Wikipedia. The platform, regardless of language, adheres to consistent core principles and structures. This uniformity allows for a more direct comparison, minimizing disparities that might arise from unique editorial policies and practices that might be present on country-specific platforms.
Quantitative data (Table 1) indicates that the English edition has the highest number of article edits, almost double that of the Korean edition and four times that of the Chinese edition, indicating more frequent revisions over time. Furthermore, its talk page is significantly more vibrant, with activity levels almost 20 times that of the Chinese edition and close to 19 times that of the Korean edition. In contrast, discussions rarely take place in both Chinese and Korean editions. This suggests that the English edition has seen more extensive discussions or debates regarding the content. The English Wikipedia edition is arguably the most extensive and active, leading in both article edits and talk page interactions. Although each edition seems to display a preference for references in its own language, the English edition distinguishes itself by being the most eclectic, integrating a broader spectrum of perspectives.

Our analysis begins with the lead sections and categorizations of the Balhae article in the three editions. Positioned at the beginning of an article, the lead provides a concise overview of central themes, capturing how each language edition encapsulates the history of Balhae. Through Wikipedia’s categorization, we explore how users from different linguistic backgrounds take a stance on the historical affiliation of Balhae. In addition to the categories of the article, we also consider the parent categories of the Balhae category. In other words, Balhae belongs to different categories as an article and as a category itself. In most instances, article categories do not deviate significantly from parent categories. However, we differentiate between these two categorizations because parent categories often include vital information that is absent in article categories. Using insights from the lead sections and categories, we then undertake a content analysis to further understand the framing of Balhae in each language edition.

Quantitative data (Table 1) indicates that the English edition has the highest number of article edits, almost double that of the Korean edition and four times that of the Chinese edition, indicating more frequent revisions over time. Furthermore, its talk page is significantly more vibrant, with activity levels almost 20 times that of the Chinese edition and close to 19 times that of the Korean edition. In contrast, discussions rarely take place in both Chinese and Korean editions. This suggests that the English edition has seen more extensive discussions or debates regarding the content. The English Wikipedia edition is arguably the most extensive and active, leading in both article edits and talk page interactions. Although each edition seems to display a preference for references in its own language, the English edition distinguishes itself by being the most eclectic, integrating a broader spectrum of perspectives.

Our analysis begins with the lead sections and categorizations of the Balhae article in the three editions. Positioned at the beginning of an article, the lead provides a concise overview of central themes, capturing how each language edition encapsulates the history of Balhae. Through Wikipedia’s categorization, we explore how users from different linguistic backgrounds take a stance on the historical affiliation of Balhae. In addition to the categories of the article, we also consider the parent categories of the Balhae category. In other words, Balhae belongs to different categories as an article and as a category itself. In most instances, article categories do not deviate significantly from parent categories. However, we differentiate between these two categorizations because parent categories often include vital information that is absent in article categories. Using insights from the lead sections and categories, we then undertake a content analysis to further understand the framing of Balhae in each language edition.

渤海国 ‘in Chinese Wikipedia

Bohai was a monarchy and multi-ethnic kingdom of the Mohe tribe dominated by the Sumo tribe from 698 to 926 AD. The founder of the country was Da Zuorong. The predecessor of Balhae State was “Koguryo,” “Sushen,” “Yilou,” “Woju,” “Wuji,” and “Mohe” in various periods in Northeast China, northern Korean Peninsula, and southeastern Russia and many other local ethnic tribes or regimes.4

The opening sentence of the lead section characterizes Balhae as a ‘multiethnic kingdom of the Mohe tribe,’ accentuating the Mohe tribe’s prominence as the primary ethnic
group within the kingdom. The Mohe were Tungusic people who inhabited present-day Northeast China. Subsequent lines suggest that Balhae’s predecessors encompass Goguryeo, Sushen, Yilou, Woju, Wuji, and Mohe, among others.

In the 10th century, the Bohai Kingdom was destroyed by Abaoji, the ancestor of the Khitan Liao Dynasty, and the old place was established as a vassal of the Dongdan Kingdom. However, the people of Bohai kept fleeing, and finally merged into the Jurchen ruling group, which was also a descendant of Mohe after the rise of the Jin Dynasty.4

The ensuing paragraph mentions that after the Khitan Liao dynasty obliterated the kingdom, the Bohai people fled and ultimately merged into the Jurchen ruling group. This statement implies that the Chinese edition of the article situates Balhae more within the context of Manchurian history than Korean history.

The Balhae article is categorized under Korean Peninsula Dynasties, Historical Regimes in Northeast China, Defunct East Asian Monarchies, and Former Chinese Vassal States. This categorization demonstrates that Chinese Wikipedians perceive Balhae as a component of both Korean and Chinese history. However, it is evident that they regard the kingdom as a peripheral aspect of their historical narrative, since Balhae is not included in the ‘History of China’ timeline.5

It should be noted that the Chinese Wikipedia has been blocked in mainland China since 2015.6 Many of its contributors thus come from Taiwan, Hong Kong, and the broader Chinese diaspora. Even before the nationwide ban, server logs from 2009–2013 indicated that only 20.7% of edits came from mainland China, compared to 39.1% from Taiwan and 25.0% from Hong Kong.7 This diversity of contributors challenges the notion that Chinese Wikipedia mirrors a single country’s viewpoint. A more monolithic perspective of the People’s Republic of China (PRC) can be found on Baidu Baike, a collaborative encyclopedia that aligns its content with the PRC’s censorship policies.

In the case of Balhae, Baidu Baike explicitly refers the kingdom as ‘a minority political power in ancient Chinese history.’8 While the PRC state has not made any official claims over Balhae’s history, most modern Chinese historians view it as an important part of Manchuria’s past – and thus, by extension, ‘Chinese history’ in the broad sense. On the other hand, the Chinese Wikipedia maintains a more circumspect stance. It also frames Balhae as a prefecture and local administrative agency of the Tang Dynasty, yet avoids claiming that the kingdom belongs exclusively to the Chinese historical narrative.

발해 in Korean Wikipedia

Balhae was a kingdom founded in 698 by Dae Joyeong, who advocated Goguryeo’s succession after the fall of Goguryeo, and existed in northern Silla and Yeonhae-ju (Primorsky) until 926.9

The Korean Wikipedia’s lead section on Balhae omits any mention of the kingdom’s multietnic character. Rather, the narrative foregrounds a general, who claimed Goguryeo’s succession, as the kingdom’s founder. While acknowledging the founder’s purported Goguryeo lineage, the article remains silent on Balhae’s affiliations with other ethnic groups. Unlike other language versions, the Korean Wikipedia delineates Balhae’s territorial extent as stretching merely to Silla (the southern kingdom of the Korean Peninsula)
and Yeonhae-ju (present-day Primorsky, Russia), thereby omitting its presence in northeastern China. Hube (2017) characterizes this type of omission as an ‘implicit’ bias, a piece of information worth mentioning but missing.

The categorization of the Korean article is confined to ‘Rulers of Balhae,’ ‘Balhae,’ and ‘History of Korea.’ Given that the first two categories are self-referential, the sole substantive category to which the article belongs is ‘History of Korea.’ This categorization inadvertently conveys the notion that Balhae is an exclusively Korean historical domain. Intriguingly, the article features a template titled ‘History of Manchuria’ – akin to other language versions – and the parent categories include it as well. Nonetheless, Korean Wikipedians opted not to incorporate the Balhae article under that category.

Within the article, the Korean entry describes various perspectives on Balhae from historical polities of differing eras – such as Qing dynasty, People’s Republic of China, Mongolia, etc. The article states that China’s recent emphasis on Balhae stems from its post-1949 ‘unified multi-ethnic national theory,’ which promotes a multi-ethnic Chinese identity over Han-centrism. The Korean article cautiously suggests that the recent revision of Balhae history in China carries political undertones, potentially undermining the integrity of Korea’s ancient historical identity in the face of perceived Chinese expansionism.

‘Balhae’ in English Wikipedia

Balhae, also rendered as Bohai, was a multi-ethnic kingdom whose land extends to what is today Northeast China, the Korean Peninsula, and the Russian Far East. It was established in 698 by Dae Joyeong (Da Zuorong) and originally known as the Kingdom of Jin (Zhen) until 713 when its name was changed to Balhae.10

The English Wikipedia presents Balhae as a multi-ethnic kingdom, refraining from emphasizing the dominance of a single tribe. In comparison to the two aforementioned excerpts, the lead section of the English Wikipedia concentrates more on factual aspects of history, thus excluding descriptions that might entail divergent interpretations. In other words, this account of Balhae has thus far proven acceptable to a majority of Wikipedians from diverse backgrounds.

The history surrounding the origin of the state, its ethnic composition, the modern cultural affiliation of the ruling dynasty, the reading of their names, and its borders are the subject of a historiographical dispute between Korea, China, and Russia. Historical sources from both China and Korea have described Balhae’s founder, Dae Joyeong, as related to the Mohe people and Goguryeo.10

All potential subjects of dispute are addressed in the subsequent paragraph, which underscores the historical intricacy of the kingdom. This portion distinctly sets the English version apart from its counterparts, as it lays out the controversial nature of Balhae in the article’s introduction. It is noteworthy that while the article categories encompass the history of China and Korea, they leave out Russia. The category ‘Primorsky Krai’ is incorporated only in the parent categories and remains absent from the article’s external links.

Compared to other language versions, the English Wikipedia forthrightly acknowledges the potential disputes regarding Balhae’s origin, ethnic makeup, and territorial
boundaries, paving the way for an open and transparent exploration of these contested historical subjects. The separate ‘Balhae controversies’ entry is dedicated to unpacking the contentious issues. In essence, the English article adopts a more encyclopedic tone, aligning closely with Wikipedia’s mission of providing information without imposing a certain perspective.

The talk pages of English Wikipedia

One unique feature of Wikipedia is that the main content page, often referred to as the ‘article,’ is simply one facet of an entry; accompanying this main page are talk pages and history pages that archive all previous versions of the article (Tkacz, 2014). While the main article page is the most visible part, it is essential to explore the associated talk pages to better grasp the dynamics of an entry. As Shirky (2009) suggests, a Wikipedia entry is better understood as an evolving process rather than a static product. Collaborative writing, by its nature, externalizes processes that remain hidden in single-authored works – in terms of ‘how the text should be written and what exactly it should contain’ (Ferschke et al., 2012, p. 777). It is within these discussions, particularly on the talk pages, that we find our primary interest.

Guided by the digital conversation analysis framework of Giles et al. (2015), we inspected the dynamics of these discussions. This approach emphasizes the importance of dissecting individual discussion threads, considering the participants, their contributions, and the sequence of the conversation (Giles et al., 2015, p. 50). Gredel further highlights the suitability of this method for Wikipedia’s talk pages, characterizing them as hypertexts that are ‘non-sequential, interactive, and dynamic’ (2017, p. 104).

Employing this methodology on the talk pages enabled us to focus on two salient discourses: the 2006 title controversy and the 2011 debate on the kingdom’s ethnic identity. These discussions encapsulate the crux of the topic’s contention, presenting a case richly threaded with strategic and communicative actions.

The profile of the top 20 editors for the Balhae entry (Table 2) reveals a diverse mix of backgrounds, encompassing various nationalities and ethnicities. A significant portion of these editors also engage with non-English Wikipedia editions, highlighting the cross-cultural dimensions of the topic (Hale, 2014). Interestingly, some editors have disproportionately focused on the Balhae entry, making a large number of edits relative to their overall contributions on English Wikipedia. This could hint at a specialized interest or even strategic motivations. The contentious nature of the Balhae topic is evident, with several editors facing indefinite blocks due to violations like sock puppetry – using deceptive online identities – or engaging in edit wars. Whereas Wikipedia is an open community, those who act against the shared ‘aims, values, and norms’ are filtered out (Pentzold, 2011, p. 712). Furthermore, the prominence of talk page edits for some users, in comparison to their article edits, indicates a preference for dialogue and debate over direct content modification.

Title controversy: Balhae or Bohai?

The title of a Wikipedia entry serves as more than a mere identifier – it is a framing device that sets the tone and context for the ensuing discourse. It is intrinsically tied


<table>
<thead>
<tr>
<th>Username</th>
<th>Description</th>
<th>Article Edits</th>
<th>Talk Page Edits</th>
<th>Total EWP Edits(^a)</th>
<th>Period of Activity</th>
<th>Note</th>
<th>Non-English WP(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qiushufang</td>
<td>Chinese History Enthusiast</td>
<td>281</td>
<td>15</td>
<td>25,060</td>
<td>2021.04–Present</td>
<td>Focused on coherence edit</td>
<td>Mongolian, Russian</td>
</tr>
<tr>
<td>Anchuhu</td>
<td>Pro-Manchurian</td>
<td>182</td>
<td>4</td>
<td>438</td>
<td>2019.02</td>
<td>Blocked (Sock puppet)</td>
<td>Russian</td>
</tr>
<tr>
<td>Koraskadi</td>
<td>Pro-Korean</td>
<td>160</td>
<td>37</td>
<td>646</td>
<td>2017.08–2020.02.23</td>
<td>Blocked (Sock puppet)</td>
<td></td>
</tr>
<tr>
<td>Historiographer (Kor Ph)</td>
<td>Native Korean; Anti-Korean Nationalists</td>
<td>146</td>
<td>1</td>
<td>14,492</td>
<td>2007.02–2011.08</td>
<td>Username changed to ‘Kor Ph’ Korean</td>
<td>Korean</td>
</tr>
<tr>
<td>Esinybro</td>
<td>Native Chinese</td>
<td>97</td>
<td>5</td>
<td>8,627</td>
<td>2019.07–2023.01</td>
<td>Chinese</td>
<td></td>
</tr>
<tr>
<td>Ulianurfnova</td>
<td>Native Russian; Anti-Korean</td>
<td>97</td>
<td>13</td>
<td>223</td>
<td>2021.11–2022.02</td>
<td>Russian</td>
<td></td>
</tr>
<tr>
<td>Zanhe</td>
<td>Native Chinese</td>
<td>71</td>
<td>0</td>
<td>138,685</td>
<td>2015.08–2019.11</td>
<td></td>
<td>Chinese</td>
</tr>
<tr>
<td>Yuje</td>
<td>Cantonese-American</td>
<td>63</td>
<td>32</td>
<td>4,459</td>
<td>2005.03–2007.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appleby</td>
<td>Pro-Korean</td>
<td>62</td>
<td>49</td>
<td>7,435</td>
<td>2005.11–2006.07</td>
<td>Blocked (Sock puppet)</td>
<td>Russian</td>
</tr>
<tr>
<td>Hatchiko</td>
<td>Native Russian; Anti-Korean</td>
<td>61</td>
<td>10</td>
<td>259</td>
<td>2019.01–2019.02</td>
<td>Blocked (Sock puppet)</td>
<td></td>
</tr>
<tr>
<td>Deiaemeth</td>
<td>Korean-American</td>
<td>55</td>
<td>109</td>
<td>3,876</td>
<td>2006.01–2007.05</td>
<td></td>
<td>Russian</td>
</tr>
<tr>
<td>IP USER 1</td>
<td>IP Address from USA</td>
<td>55</td>
<td>24</td>
<td>55</td>
<td>2023.02</td>
<td>Made 55 edits during 2 h</td>
<td></td>
</tr>
<tr>
<td>Yprpyqp</td>
<td>Interested in Chinese History</td>
<td>52</td>
<td>0</td>
<td>15,652</td>
<td>2017.05–2017.07</td>
<td>Focused on grammar edit</td>
<td></td>
</tr>
<tr>
<td>Kuebie</td>
<td>Native Korean</td>
<td>45</td>
<td>4</td>
<td>1,621</td>
<td>2007.07–2010.11</td>
<td>Blocked (Edit-warring)</td>
<td></td>
</tr>
<tr>
<td>Jagello</td>
<td>Pro-Korean</td>
<td>41</td>
<td>13</td>
<td>479</td>
<td>2011.12–2019.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IP USER 2</td>
<td>IP Address from USA</td>
<td>37</td>
<td>8</td>
<td>79</td>
<td>2007.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kaustritten</td>
<td>Native Russian; Anti-Korean</td>
<td>34</td>
<td>16</td>
<td>85</td>
<td>2020.10–2021.06</td>
<td>Blocked (Sock puppet)</td>
<td>Russian</td>
</tr>
<tr>
<td>Ni</td>
<td>Native Chinese</td>
<td>30</td>
<td>20</td>
<td>166,420</td>
<td>2005.10–2012.12</td>
<td>WP Administrator</td>
<td></td>
</tr>
<tr>
<td>Endroit</td>
<td>American; Anti-Korean</td>
<td>28</td>
<td>34</td>
<td>11,217</td>
<td>2006.03–2007.09</td>
<td></td>
<td>Chinese, Japanese</td>
</tr>
</tbody>
</table>

Note: The information of each user is mostly obtained from their user page.
\(^a\)The number of total edits made on the English Wikipedia by a user.
\(^b\)The non-English Wikipedia editions edited by a user.
to issues of identity and recognition, often becoming a symbol of validation or denial of a particular perspective (Góngora-Goloubintseff, 2020; Gredel, 2017). Furthermore, the act of titling is a manifestation of power dynamics, with the authority to name and define subjects often leading to struggles over control of knowledge representation (Ford et al., 2015). The NPOV policy is frequently contested in the process of title selection, as perceived neutrality can vary greatly among individuals (König, 2013; Matei & Dobrescu, 2011; Rogers & Sendijarevic, 2012). Additionally, the translation of titles in a multilingual environment like Wikipedia can introduce additional layers of controversy due to cultural, political, or ideological biases inherent in language (Jones, 2018; Kumar, 2017).

Building upon this complexity, the lack of information regarding the languages spoken by the Balhae people adds another layer of difficulty to the issue of how to title the entry. In line with most East Asian kingdoms of the time, Balhae used an official writing system based on classical Chinese characters. Nevertheless, historical accounts have not detailed how the Balhae people pronounced these characters. Modern Chinese and Koreans read ‘渤海’ as Bohai and Balhae, respectively; however, neither of these readings accurately represents the pronunciation at the time (Crossley, 2016). Employing either name inevitably entails a degree of modern political projection.

The title of the Balhae article has been a point of contention since its inception. Hence, the edit history and the talk pages of the entry display numerous strategic actions. The article initially bore the title ‘Bohai’ for the first three years, during which time several attempts were made to change it to ‘Balhae.’ The debate over the article’s title reached a critical juncture in 2006 when a user Deiaemeth took the initiative to move the content of the ‘Bohai’ article to a new page titled ‘Balhae.’ This action sparked a vigorous discussion among the community, bringing forth the complexities of naming and representation in a global platform.

Nanshu, the first editor of the article, proposed a poll to determine whether Chinese pinyin should be adopted as the default romanization. The vote for using pinyin sparked a two-week-long discussion that morphed into a debate over the ownership of history. Nanshu’s argument was rooted in the premise that the primary sources on Balhae were in classical Chinese, and that ‘Bohai’ is the term predominantly used in Chinese and Russian academia. Nanshu further dismissed English sources as lacking authority:

Appleby seems a worshipper of Britannica, but actually, Britannica and other English sources are not authoritative for this field because, as far as I know, there is no comprehensive work on Bohai in English.11 (Nanshu, 10:24, 21 February 2006)

However, this viewpoint was met with resistance from users Appleby and Deiaemeth, who advocated for the use of the Korean romanization (Balhae, Parhae, Palhae, etc.) in line with Anglophone academia and publications. These users invoked Wikipedia’s verifiability policy, which prioritizes English-language sources to ensure that readers can readily verify the accuracy of the information presented. This policy, they argued, underscores the need for English Wikipedia to reflect the conventions of Anglophone scholarship, regardless of whether English sources are more or less authoritative than the non-English ones:
I’m just not sure why we need to do this convoluted analysis and original research of foreign language sources, when Wikipedia is an English language encyclopedia and its policy is to use common English names and to rely on reputable unbiased sources.\(^{11}\) (Appleby, 17:47, 21 February 2006)

Advocates for the title ‘Bohai’ resisted the alternative ‘Balhae,’ arguing that the latter carried an inherent Korean bias. They posited that ‘Bohai’ was a more neutral choice, given its common usage in China and Russia – countries with direct access to and research capabilities concerning the archaeological sites in question. Conversely, proponents of ‘Balhae’ compiled a list of English-language sources that referred to the kingdom as ‘Balhae,’ challenging the ‘Bohai’ advocates to cite reputable English sources that used the Chinese reading.

The Encyclopaedia Britannica entry, which employs the Korean romanization ‘Parhae,’ became a significant point of contention. The ‘Bohai’ advocates invoked a passage from the Verifiability Policy, which warns against the potential unreliability of unsigned entries written by freelancers with limited expertise. They argued that original texts written in the 10th – 11th centuries, such as the \textit{Old Book of Tang} and the \textit{New Book of Tang}, should take precedence over sources like the Encyclopaedia Britannica:

\begin{quote}
Appleby cannot be allowed to make edits which blatantly contradict information on the authoritative primary sources such as the “Old Book of Tang” (945) and the “New Book of Tang” (1061).\(^{11}\) (Endroit, 00:30, 6 March 2006)
\end{quote}

However, this argument was countered by the assertion that the cited passage is from a non-binding \textit{guideline}, not a non-negotiable \textit{policy}. The ‘Balhae’ advocates maintained that ancient, non-English historical texts are often subject to interpretation and original research, thus complicating their use as definitive sources. Similar to typical controversial entries, a lengthy discussion continued and became repetitive; both sides repeatedly cited the same justifications and regulations to push forward their arguments, which exemplify strategic actions (Schneider et al., 2010).

On 28 February 2006, a shift occurred in the ongoing debate. The ‘Balhae’ advocates withdrew their votes and the poll, citing Wikipedia’s policies and guidelines that discourage the use of polls as a means to resolve disputes. While the ‘Bohai’ advocates outnumbered the ‘Balhae’ counterparts, the broader user community accepted this decision, recognizing that the purpose of the poll was to establish consensus rather than to make a final decision. This marked the end of the extensive discussion, and the title ‘Balhae’ has remained in place since then. While there have been sporadic attempts to revert the title to ‘Bohai,’ these efforts have not sparked any substantial discourse.

The interactions between the users were marked by efforts to justify their respective preferences, drawing on Wikipedia policies, historical references, and linguistic arguments. The majority of comments displayed the characteristics of strategic actions. The manner in which both sides referenced policies and guidelines seemed more intent on discrediting the opposition than seeking common ground. Admittedly, both sides entered the discussion with distinct goals. Through the course of discussions, this tug-of-war ultimately resulted in a more neutral development of the article. While direct consensus was not achieved, the lengthy discussion implicitly underscored that the exclusive use of a specific romanization contradicted NPOV.
As a result, the article now presents both Korean and Chinese readings side-by-side, with the order varying according to context. Notably, the ‘Balhae’ advocates modified their stance over time, shifting from portraying Balhae as a Korean state to acknowledging its multi-ethnic character:

To remove a distraction, I will remove the wording “Korean kingdom” from the intro, although I think it’s a fair description of what the references say.\textsuperscript{11} (Appleby, 08:11, 27 February 2006)

I think this will placate user:Ran and user:Yuje, as they were upset regarding terming Balhae exclusively as a Korean state. While I regard Balhae does have significant place on Korean history, various ethnic groups also made up Balhae and that should be not disregarded. As for the content, I hope we can work something out.\textsuperscript{11} (Deiaemeth, 08:32, 27 February 2006)

Despite their adjustments, it is clear that the foundational beliefs of these users remained unchanged. Their concessions seemed more about negotiating neutrality than collaboratively achieving it. This suggests that ‘collective’ might be a more apt descriptor for Wikipedia’s knowledge-building process than ‘collaborative.’

Drawing on the observation from Ford et al. (2015) that those adept at the complex policies of Wikipedia often prevail over those with disciplinary knowledge of the topic at hand, we see a similar dynamic at play in the Balhae title debate. While the ‘Bohai’ advocates presented a lot of predominantly non-English sources, the ‘Balhae’ advocates, though fewer in number, were able to prevail by strategically citing the verifiability policy that favors English sources. The resolution of the debate in favor of the ‘Balhae’ advocates was less a reflection of the historical validity of their viewpoint than an outcome of their better understanding of Wikipedia’s policies and norms of interaction (Bender et al., 2011).

**Ethnic identity debate: to whom does Balhae belong?**

The discourse surrounding the Balhae entry on Wikipedia has frequently veered into questions of ethnic identity, prompting contributors to question the historical characterization of the kingdom. In particular, instances of vandalism often involve the insertion of nationalities, such as labeling Balhae as a ‘Korean’ or ‘Chinese’ kingdom, into the lead section without discussion. As an alternative, User Cydevil first proposed the use of terms ‘proto-Koreans’ and ‘proto-Manchus’ to describe the ethnic composition of the kingdom, and in turn suggested ‘multiethnic’ as a simpler and more neutral modifier.

This shift in discourse is significant. It marks a transition from a binary portrayal of Balhae to a more neutral presentation of the kingdom as a multiethnic entity. The term ‘multiethnic,’ despite its inherent ambiguity, mitigates potential tensions and fosters resolution. Whereas pro-Chinese users emphasized the presence of Mohe tribe and pro-Korean users highlighted the refugees of Goguryeo, the bottom line was that both factions could agree, at the very least, Balhae consisted of various ethnicities. This approach aligns with Matei and Dobrescu’s concept of ‘settling conflict through ambiguity,’ wherein the systemic ambiguity of the NPOV policy serves to alleviate conflict (2011, p. 41). As of 2023, sporadic acts of vandalism continue, with attempts to replace the term ‘multiethnic’ with ‘Korean’ or ‘Chinese,’ but such edits are promptly reverted.

The development of the entry also drew the attention of contributors who did not align with either side. However, these contributors initially faced resistance. For instance,
User Huang Tai Ji posited that Balhae represents a ‘shared history’ between contemporary Korea and the People’s Republic of China, embodying the history of Manchuria rather than that of either nation. Despite Huang Tai Ji’s comment aligning with the NPOV policy, User Korsentry challenged this assertion, mockingly proposing to invite Russia to share Balhae’s history as well. This resistance underscores the reluctance of some contributors to accept the concept of a shared history, maintaining that the kingdom’s past was exclusively linked to one country or the other.

By the early 2010s, discussions concerning the ethnic identity of Balhae began to adopt a more moderate and detached approach. User Ecthelion83 highlighted the inherent challenge in applying modern ideas to a historical entity that did not survive into the contemporary era. Ecthelion83 emphasized the uncertainty surrounding Balhae’s self-identification with any specific ethnicity or culture, given the limited availability of primary sources:

I guess the major problem with categorizing Balhae, an ancient kingdom, using (relatively) modern ethnopolitical/ethnocultural definitions has greatly to do with the kingdom’s own self-identification … we do not even know how strongly the people of Balhae actually self-identified to any ethnicity or culture, since the only descriptions of this population in ethnic terms come from sources outside of Balhae itself.12 (Ecthelion83, 10:38, 12 March 2011)

As Rosenzweig (2006, p. 122) characterizes ‘neutrality’ as a ‘founding myth’ for Wikipedia, the NPOV policy provides a guiding framework for users to scrutinize and recalibrate their arguments. While some contributors initially intended to advance their POVs, aligning their arguments with the NPOV policy often leads to a deeper introspection of entrenched nationalist narratives. This leads to an overall refinement of article content and a discourse depth seldom observed in other digital platforms.

This critical observation underscores the evolving dynamics of the debate, as users increasingly acknowledge the ideological impasse and recognize the importance of disentangling modern political circumstances from historical records. This shift led to the creation of the ‘Balhae Controversies’ entry when user Koraskadi transferred the ‘politicization’ section of the ‘Balhae’ article to a separate entry, which will be further elaborated in the discussion.

**Discussion**

The development of the Balhae entry on English Wikipedia presents an insightful lens into online knowledge production and historical narratives in a globalized context. Central to this dynamic are memory, heritage, and cultural ownership. While the Balhae entry might seem like an isolated issue, it represents a broader struggle: the tug-of-war between different nationalistic ideologies seeking validation on a global platform.

In this regard, the democratic structure of Wikipedia is both a strength and vulnerability, as the platform invites instrumental and strategic actions. The platform’s decision to allow edits from non-registered users is rooted in its commitment to harnessing the wisdom of the crowd. While instrumental actions can sometimes introduce biases or conflicts, they can also serve as catalysts for discussion, prompting the community to re-evaluate content, sources, and perspectives. That is to say, Wikipedia’s allowance for instrumental actions and non-registered edits is a deliberate choice to prioritize
openness and inclusivity, while relying on its community and built-in mechanisms to manage and mitigate potential challenges.

Strategic actions are evident in the talk pages due to the overlapping nature of strategic and communicative actions. The interactions at the talk pages showcase that adherence to the policies do not necessarily make interactions more communicative. However, Wikipedia’s policies and community mechanisms guide these actions towards a more neutral outcome, since it is often simpler to guide an edit towards neutrality than to change a user’s inherent perspective.

Although initial contributions might be motivated by personal or strategic inclinations, the collective discourse, shaped by the platform’s structure and principles, steers towards rationality (Hansen et al., 2009). The Balhae entry bore witness to ostensibly strategic actions – edits motivated by seemingly nationalistic drives, with aspirations to dominate the overarching story. This is markedly evident in attempts to brand Balhae either as distinctly ‘Korean’ or ‘Chinese.’

Yet, as Wikipedia’s user base evolved and the NPOV principle solidified, the discourse gradually became to display more communicative actions. The narrative shift from entrenched nationalist tags to more inclusive descriptors like ‘multiethnic’ underscores this development. Instead of pushing individualistic views, users began to collectively seek a more holistic understanding of Balhae’s history. Though not without turbulence, this shift mirrors Habermas’s scenario of ideal speech, characterized by genuineness, legitimacy, and clarity.

The division of the article into ‘Balhae’ and ‘Balhae controversies’ also marked a pivotal moment in the discourse. This separation acknowledges the existence of multiple, valid interpretations of history, while also maintaining a clear distinction between established historical facts (as discussed in the talk pages) and their interpretation. When recurring topics arise, particularly for new readers of the article, the active contributors to the Balhae entry direct them to the Balhae controversies entry to continue their discussion there. This approach could serve as a model for handling other contentious historical topics on similar online environments.

**Conclusion**

This case study illustrates the potential of Wikipedia as a platform where individuals with diverse, even conflicting perspectives can collectively craft an article (Jirschitzka et al., 2017) that overcomes deep-seated historical convictions over time. By looking at the development of the Balhae article on the English Wikipedia, we have explored the intricacies of historical discourse in the digital age.

However, the study’s focus on a singular entry entails limited applicability. The insights from the Balhae article may not necessarily apply to other contentious historical topics on Wikipedia, given how distinct each controversy is.

Future research could broaden this scope by examining other contentious articles, providing a more holistic view of historical representation on online spaces. Exploring contemporary nationalistic disagreements – particularly territorial or colonial histories among East Asian nations – could be enlightening. After all, Wikipedia articles on the history of China and Japan are known to pose large interlingual disagreement (Samoiilenko et al., 2017, p. 216). Given that, the dyad and triad interactions among Wikipedians
from South Korea, China, and Japan might offer insights into how national narratives interplay at the grassroots level.

Further, understanding the motivations of top contributors to contentious historical articles can provide a clearer picture of potential biases. Comparing prolific cross-topic contributors versus those focused on history or contentious subjects could reveal some interesting patterns too.

Outside of Wikipedia’s confines, it would be worth looking at how external political developments or diplomatic strains between nations impact related Wikipedia edits and discussions. A time-series analysis comparing Wikipedia with multilingual academia might yield intriguing findings as well. Additionally, the influence of Wikipedia on ‘public history’ might be gauged – surveys or experiments could be conducted to discern how Wikipedia’s portrayal of historical events shapes public knowledge and perceptions (De Groot, 2009).

The English Wikipedia represents an uncharted domain within digital communication studies in the East Asian context, requiring meticulous investigation of the underlying social and cultural processes at play. It is our hope that this article sparks a robust dialogue about the clashes between monolithic, nationalist historical narratives. The potential avenues for future research are wide-open, with plenty of unanswered questions still awaiting scholarly inquiry.

Notes
1. The name of the kingdom is rendered 渤海 (pinyin: Bóhǎi) in contemporary standard Chinese, 발해 (romanized: Balhae) in Korean, and Бокхай (romanized: Bokhay) in Russian; in this study, we have opted to use the name Balhae, following the title employed in the Wikipedia article and to avoid lexical disambiguation.
2. https://xtools.wmcloud.org
8. https://baike.baidu.com/item/渤海国/1551268
11. https://en.wikipedia.org/wiki/Talk:Balhae/Archive_1

Disclosure statement
No potential conflict of interest was reported by the author(s).

Funding
This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the metaverse support program (IITP-2024-RS-2022-00156435) grant funded by the Ministry of Science and ICT (MSIT); Basic Science Research Program through
the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2023-00241749).

Notes on contributors

Jonghyun Jee is a master student at the Graduate School of Metaverse, Korea Advanced Institute of Science and Technology (KAIST). His research interests include digital nationalism and computational social sciences.

Byungjun Kim, Ph.D. (Sungkyunkwan University), is a research assistant professor at the Center for Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology (KAIST). His research interests include digital humanities, natural language processing, and computational social science.

Bong Gwan Jun, Ph.D. (Seoul National University), serves as a professor at the School of Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology (KAIST). His scholarly endeavors focus on digital humanities, digital storytelling, and Korean literature.

References


