

## **Connective Action and Affective Language: Computational Text Analysis of Facebook Comments on Social Movements in South Korea**

SHIN HAENG LEE  
Sejong University, South Korea

TAE YUN LIM  
Hongik University, South Korea

The present study addresses the question of how the 2016 “candlelight vigil” in South Korea developed through digitally networked activism in opposition to then president Park. Specifically, we examine whether the anti-Park group and its countering pro-Park group manifested different semantic structures in their comments on Facebook pages. To do so, we collected 144,639 and 29,133 comments on the anti-Park and pro-Park pages, respectively, over 5 months (October 2016–March 2017) and constructed semantic networks from the comments made during the heady days of activism. Our analysis of word co-occurrence and community structure for each network demonstrates that the anti-Park activism engaged in random and polysemic patterns of word use. In contrast, the organization-based pro-Park activism exhibited scale-free and semantically coherent organizing. This finding explicates the linguistic mechanism by which the connective logic of engagement affords affective expressions that generate loosely formatted but collaboratively coordinated identities to intensify feelings of engagement.

*Keywords: connective action, affective language, Facebook activism, semantic network, community detection, computational text analysis*

For the time being and in the moment, they are our means for feeling our way into worlds we cannot experience directly, and as such, they mean something. (Papacharissi, 2016, p. 32)

When scholars have considered contemporary social movements in East Asian democracy such as the 2014 Occupy movements in Hong Kong and Taiwan, as well as the 2016 “candlelight vigil” against then

---

Shin Haeng Lee: shinlee@sejong.ac.kr

Tae Yun Lim: tyim27@hongik.ac.kr

Date submitted: 2018–10–10

<sup>1</sup> This work was supported by a grant from Sejong University in 2018 (Grant #20180412).

Copyright © 2019 (Shin Haeng Lee and Tae Yun Lim). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <http://ijoc.org>.

president Park in South Korea (hereafter, Korea), the role of digital network technology in mobilization has been a contentious subject (Chen, Ping, & Chen, 2015; F. L. F. Lee & Chan, 2016; S. H. Lee & Fu, 2019). Some researchers argue that social media reshape the way in which people engage in public life across the region, and therefore contemporary political activism enables the traditionally marginalized (Howard, 2015; S. H. Lee & Fu, 2019). However, this cyber-optimistic view invites critical arguments in line with “net delusion” (Morozov, 2011). Even if the technology is viewed as an easy way to gain access to grassroots movements or collective organizing at an unprecedentedly low cost, it is naïve to underestimate the agency of preexisting institutions, social structures, and cultures in maintaining the status quo (Diamond, 2010). More importantly, discussion of whether digital technology democratizes citizen participation is inconclusive unless we can understand what exactly is enabled by the technology in the mass mobilization (Howard, 2015).

To fill this gap, this study demonstrates how social media serve as a virtual venue for large-scale mobilization that accommodates personal expressions about activism and forms an individual and collaborative identity about the cause in question. Of course, it is not surprising that Facebook pages provide activists and their groups with a means of framing what is going on in public life and engaging the mass citizenry. But what is notable in digitally networked activism is that in the absence of formal organizations or shared ideologies, social media engage individual citizens around a certain political event (Bimber, Flanagin, & Stohl, 2012). Activism discourse builds on personal frames of narratives that are not easily consolidated into a collective voice but can be loosely coordinated for activism with a certain degree of modularity. In doing so, out of collaboratively coded expressions on the platform there emerges flexibly a political identity, which in turn enables the participants’ contribution to discourse building (i.e., followers commenting on organizers’ posts). This view draws on Bennett and Segerberg’s (2013) logic of connective action, which allows us to examine how personal voices coalesce into a collaborative imagination through casual engagement in activism, thus differing from conventionally organized collective action. We therefore visit the logic of connective action that organizes personal frames and coordinates individual engagement in political activism by analyzing user comments on Facebook pages created for the candlelight movement against then president Park Geun-hye (anti-Park) and its countermovement (pro-Park) in Korea.

As an extension of Bennett and Segerberg’s (2013) discussion, this study also uses computational text analysis of the language used in the comments on activism posts that were retrieved from a series of Facebook public pages associated with the anti- and pro-Park movements. This textual analysis examines how different semantic networks develop out of verbal expression and communication in the context of connective and collective logics of large-scale mobilization. Although previous studies have analyzed political groups’ word use to examine their identity in relation to political cleavages, there are few studies, to our knowledge, on semantic maps of language corpus with respect to different logics of political action working in large-scale mobilizations. To this end, the literature on social movements is first reviewed with respect to why collective identity formation occurring on social media platforms matters and how word use is manifested differently in the context of semantic networks. For the methodology, we give an account of our data collection from the anti- and pro-Park public activism pages on Facebook and our text analysis of a linguistic corpus of comments about posts on the pages.

We found that Facebook engagement in anti-Park activism involves a much more decentralized structure of performative language in that its cross-referentiality connects more diverse individual viewpoints to the collaborative identity formation. By contrast, pro-Park pages remediate the same pattern of word use in the organizing of collective discourse, so that politically established ideas or opinions are reproduced. We argue that the anti-Park activism was an affect-driven movement that connected people's diverse perspectives and negotiated more loosely formatted, but still politically active, identities. Affective language in social media make possible an individual level of performance that further connects the personalized and heterogeneous frames of enunciation to the conversations of more collectively assembled social movements.

### **Case Background**

The anti-Park movement was sparked by a special report on the scandal involving the president on October 24, 2016, carried by JTBC (a major Korean nationwide cable TV network and broadcasting company). Although Park moved quickly to push for a constitutional amendment to swing media attention, her corruption was exposed further, resulting in strong public sentiment forming to call her to account for it. Afterward, thousands of activists and even ordinary citizens started rallying on the streets near the president's house to push for her impeachment on October 29. This event became the beginning of the anti-Park movement, the so-called candlelight vigil, in which tens of thousands of people, without any political affiliations or apparent ideological commitment, went to the streets to demand the president's resignation. On December 3, more than 2 million people joined the rally held across the country, according to the organizers. The protesters had occupied downtown Seoul, the capital city, and other major cities every weekend until the president was finally discharged on March 10, 2017. But the anti-Park movement also led to a pro-Park movement to undermine public dissent. This countering group of protesters, in contrast, consisted of the ideologically conservative who identified themselves with the ideologies of Park Geun-hye's father, the former president Park Chung-hee, the most admired conservative president that Korea has had. These protesters shared a strong identity at variance not only with North Korea but also with the progressives; most participants carried the Korean national flag, Taeguekgi, which symbolized their rally.

During the candlelight movement, online publics in anti-Park activism were affectively rendered on Facebook pages from the day JTBC reported the statements about presidential scandal, on October 24, 2016. Facebook was indeed the most dynamic platform for online mobilization in South Korea at that time, because the service's market penetration outpaced Twitter and Instagram as well as other local social media in time spent on it (Nielsen KoreanClick, 2016). And over the period of the large-scale movements, some noticeable changes were found in the anti- and pro-Park activism pages on Facebook, especially in their ways of engaging users. One of the features of the anti-Park movement is that it kept populating diverse narratives and causes among the pages so that many public and private pages were created to stream a mixture of opinions, news, and general sentiment online. On the other hand, pro-Park pages put forward a coherent slogan to act on. Such technological affordances for group organizing enabled Facebook to be considered the most prominent platform where the Korean contours of social movements could be observed in the virtual setting.

## Theoretical frameworks

### *Connective Action and Digital Media*

To account for mobilization for social movements, the literature suggests that organizations help individuals bear the costs of participation "on their own" by providing selective incentives such as "material goods, financial rewards, moral feelings of altruism, [and] solidarity with desirable others" (Bimber, Flanagin, & Stohl, 2005, p. 368). Especially, sociologists highlight the effective role of formal organizations in collective-identity formation, as well as resource mobilization: Within certain organizational structures, individuals are consolidated into a collective identity so that their grievances are linked with who they are (McAdam, Tarrow, & Tilly, 2009; Polletta & Jasper, 2001). Collective identity is thus a prerequisite for social movements, as it provides people with "activists' shared definition of their situation, the expressive character of all action, the affective bonds that motivate participation, the experience of solidarity within movements, and others" (Polletta & Jasper, 2001, pp. 284–285).

The global diffusion of personal social technology has, however, enabled the emergence of digitally networked publics who build on both actual and imagined communities across national borders and develop into large-scale social movements (Howard & Hussain, 2013). Indeed, a contemporary form of political activism is related to technological affordances of social media that facilitate interpersonal relations and individual groups in organizational effects on mobilization and coordination (Bimber et al., 2012). Regarding why contemporary social movements revolve around digital media, Bennett and Segerberg (2013) put forth the logic of connective action: Digitally enabled networks lend individuals certain organizational structures to act in a shared cause. Papacharissi (2015) adds to this line of scholarship by arguing that social media, in which the "networked and condensed performance of the self" becomes possible, help users craft ideologically more diverse and polysemic identities "without compromising their own sense of who they are" (pp. 99–100). For her, self-identity of the users is constantly reformatted, displaced, and extruded by the affective language styles in social media. The social-media users thus constantly experience becoming a mutated self, without necessarily "compromising [their] own sense of self" (Papacharissi, 2015, p. 112) to the traditional sense of identity.

Of course, conventional wisdom argues that political identifications emerge from the concrete communities insofar as mobilizing identities are based on framing efforts to be culturally resonant with potential adherents and constituents (Benford & Snow, 2000). In the same vein, the participants in a protest should identify with the cause of its organizers, because a sense of solidarity and commitment comes from shared beliefs and ideologies. But in the context of connective-action theory, the process by which participation is mobilized is scaled down to individuated levels from hierarchically structured and organization-mediated ways (Bennett & Iyengar, 2008). The trend of personalized politics is indeed manifested in the micromobilization process by which individual voices are coordinated within the conversational context of communication artifacts such as trending tags and thus distinguished from the traditional process of collective identifications.

In today's digital era, political elites and stakeholders in decision making increasingly encounter the rise of social movements whose constituents are hard to attribute to a certain group identity or

ideological affiliation. Instead, the disaffected are characterized by flexible identifications—so much so that participatory citizens are increasingly endowed with diversified personal interests and amorphous ideologies rather than hierarchically structured and traditionally organized mass politics (Dalton & Wattenberg, 2000). The term “loosely tied, opt-in/opt-out networks” refers to the organizational characteristic of connective action: Personal motives and diverse issue definitions are connected to the mobilization process by which participants do not necessarily embrace the same ideology or identity for collective action (Bennett & Segerberg, 2013). Thus, what distinguishes the logic of connective action is that diverse personal expressions and performances are organized in digitally networked activism so that “individuals define issues in their own terms and . . . network with others through social media” (Bennett & Segerberg, 2013, p. 59). From this perspective, digitally networked activism enables integration and coordination of individuals’ multiple, loosely coupled identities in large-scale mobilization.

The collective identity of political groups has been often investigated in relation to their word use, which is assumed to mediate the expression and communication of “embedded values, beliefs, and social identity” (Neiman, Gonzalez, Wilkinson, Smith, & Hibbing, 2016, p. 214; see also Grimmer & Stewart, 2013; Lakoff, 2002). Through natural language processing and a lexicon-based approach to opinion mining, prior work found significant differences in discourse that reflected the personal values and ideologies of legislators who sent public messages through their public Twitter accounts (Jones et al., 2017).<sup>2</sup> This line of scholarship examines utterance observations by comparing value-laden languages that are assumed to reflect different contours of political ideologies (Laver, Benoit, & Garry, 2003). Instead, in this study, we focus on different word co-occurrence behavior performed by politically opposing groups during the time period of a large-scale social movement.

### ***Affective Language in Facebook Activism***

Regarding digitally networked activism, many people join in large-scale action taking place online during a period of political upheaval. But Facebook affords the logic of connective action that organizes mobilizing personal frames and coordinates group identities. Through Facebook pages, specifically, organizers’ postings develop into a place for personalized communication that invites comments from the page’s followers, who express spontaneous opinions that they have on the troubled situation. This digital-networking platform has the potential to reshape social relationships in civic associations through its automated algorithmic system, which encourages a multilayered network of networks by bridging latent weak-tie networks that would otherwise be unconnected (Bakshy, Messing, & Adamic, 2015; Howard, 2015). Doing so encourages feelings of connectedness and individual engagement in the organizers’ narrative without any conventional pathways to collective organizing, such as political identifications (Coretti & Pica, 2015).

---

<sup>2</sup> Of course, Neiman et al. (2016) found that the vocabulary of politicians did not systematically demonstrate differences in the respective value systems of Republicans and Democrats in the U.S., though the parties were supposed to be polarized. One possible reason for the similarity of partisan vocabularies is that, as suggested by Benford and Snow (2000), the conflicting groups compete for meaning making in their attempt to frame public issues using the same publicly accessible terms while the affective language discriminates collective identities.

Given the logic of connective action that enables digital networks to create the organization principles for any large-scale action, Bennett and Segerberg (2013) differentiated between crowd-based action networks, in which diverse individuals and platforms constitute layered networks to organize activity, and organization-based networks, in which the organizers arrange “the digital engagement mechanisms for individuals to use in their own networking and participation choices” (p. 196). The former involves technologies as agents in the creation of organization to coordinate personal-action frames and carry diverse identities; thus, “a large number of people may still join in action, but the identification is derived more through inclusive and diverse large-scale personal expression than through group or ideological affiliation” (Bennett & Segerberg, 2013, p. 25). But the latter positions conventional organizations in the center of mobilization that benefits instrumentally from digital networks as a new means of collective-identity formation; in this framework of large-scale mobilizations, the logic of action depends on “organizational solutions that focus on the importance of leadership, resources, framing, and brokerage among resource-rich organizations aiming to create incentives for individuals to join in a common cause, take risks, and participate in the networking process with others” (Bennett & Segerberg, 2013, p. 196). From this perspective, we assume that affect-driven social movements develop in the context of crowd-based action networks rather than organization-based networks.

When it comes to affect-driven social movements, Papacharissi (2015) adds that social media such as Facebook serve as conduits for networked publics who are collaboratively organized through affective expression: that is, affectively charged posts and performative language invite people’s diverse responses and connect their feelings of engagement. Especially, social media construct a more liberated yet collaborated narrative structure, the mediality of which lends a certain form to unrepresented social events or groups, so that their voices can be shaped into more collective and visible narratives (Papacharissi, 2016). She also emphasizes how these social-media platforms provide users with affectively driven yet performative linguistic agency, so that they can constantly “infuse new meaning into the texture of a performance, frequently through linguistic play” (Papacharissi, 2015, p. 98). She believes that affective and performative narratives or news comments in online space can form a more open and loosely collaborated social structure—what Raymond Williams (1977) calls a “structure of feeling.” Papacharissi’s (2015) use of the terms “performative” or “performative language” seem to be related to Austin’s (1962) language theory in *How to Do Things With Words*: He argues that “all utterances necessarily become performative” (p. 144) and have an interpellative function that constitutes one’s social identity. But, at the same time, he makes clear that not all speech functions are necessarily efficacious, and some of them, especially “perlocutionary speech acts,” bring in unintentional/unexpected meanings and thus emancipate speakers from their socially and ideologically interpellated identities (Austin, 1962, p. 144).<sup>3</sup>

Papacharissi’s (2015) emphasis on performative language in social media also has a theoretical resonance with Judith Butler’s performance theory.<sup>4</sup> Butler (1997) explores how one’s political identity, even once constituted in the dominant discourse, can be reconstituted quite differently within different linguistic arrangements. In the same vein, Papacharissi’s (2015) notions of “affective public” and “performative language” emphasize the possibility that Internet users’ identities can be equipped with multiple, distinct and even

---

<sup>3</sup> Austin (1962) divided speech acts into three categories: locutionary, illocutionary, and perlocutionary.

<sup>4</sup> Butler’s (1997) performance theory was also influenced by Austin’s notion of perlocutionary speech acts.

conflicting potentials to “reinvent personal and dominant narratives” (p. 98) and form a loosely collaborated structure of storytelling. In line with this reasoning, the following question is posed:

*RQ: Will anti-Park activism on Facebook accommodate affectively charged expressions through performative language? And will its word co-occurrence behavior allow a semantically more pluralized interpretation of the political event that is not available in collectively organized pro-Park activism?*

### **Method**

To gather data for the present study, we compiled a list of activism pages on Facebook in relation to the anti-Park movement and its countermovement that were publicly accessible on March 10, 2017, the day on which the Korean Constitutional Court upheld the National Assembly’s vote to impeach then president Park. From these pages, we obtained information about comments on each post published on the pages from October 24, 2016, the day on which JTBC started reporting the presidential scandal. In sum, our data set was composed of 172 pages, from which we downloaded 173,772 associated comments in the time span of five months. This framework of sampling seeks to capture temporal characteristics of the ephemeral web sphere of activism with validity by fixing the boundaries of Facebook pages that share content links with each other, feature activism, and invite user engagement (Schneider & Foot, 2004).

#### ***Sampling of Facebook Pages***

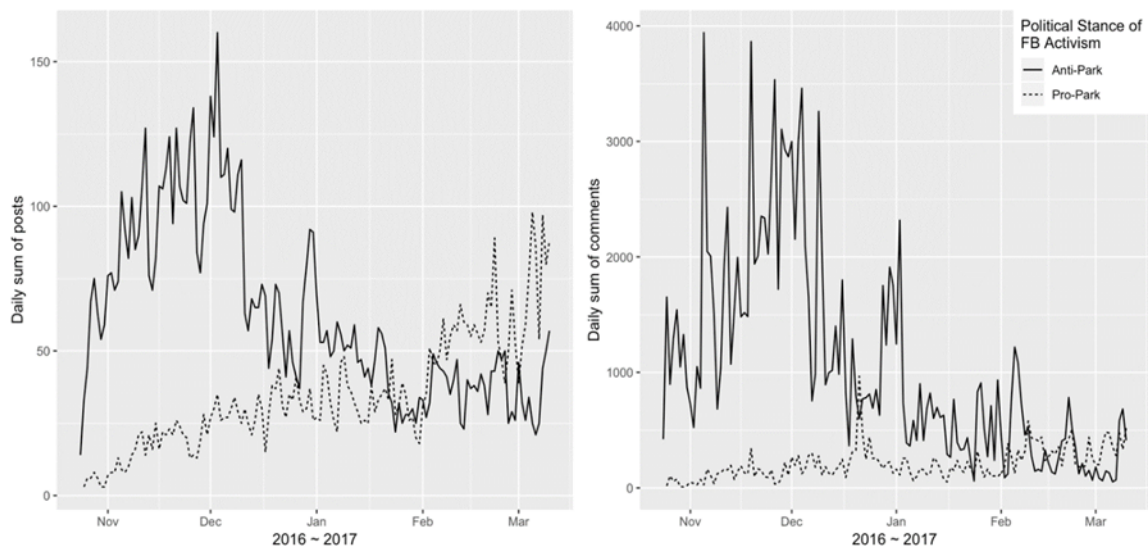
Our method for collecting the Facebook data followed several steps. First, specific issue-related terms in Korean (e.g., 박근혜 [Park, Geun-hye], 대통령 [president], 탄핵 [impeachment], 사퇴 [resignation], 촛불 [candlelight], 태극기 [Taeguekgi]) were entered into the advanced group search-box function embedded on Facebook (see Woolley, Limperos, & Oliver, 2010). Second, from the search results, we preliminarily sampled a cluster of public pages created to support or counter the anti-Park movement and labeled as community or political group.<sup>5</sup> Third, from the set of seed pages, additional pages were retrieved and archived when they shared a post. Fourth, given the purpose of this study, Facebook pages were included in the sample only if they were not apparently affiliated with or identified as political institutions or conventional organizations such as government agencies, political parties, news media, labor unions, and formal civic groups.

The above process of Facebook searching yielded 41 activism groups sorted in descending order with respect to group membership, then divided into 31 anti- and 10 pro-Park pages. These seed pages were then used to employ a purposive sampling technique to include all public Facebook group pages in relation to either anti- or pro-Park movements. As a result, a total of additional 313 pages was identified that shared a hyperlink to mobilizing content with the seed pages. From these pages, 128 anti- and three

---

<sup>5</sup> Our sampling frame is subject to potential change for two reasons: First, public pages on Facebook do not provide information about when they were created and when they no longer existed (Woolley et al., 2010). Second, the Facebook search function yields customized results relevant to user interests in line with a history of previous activities. This limitation was addressed, however, by ensuring that the same search results could be obtained through another user ID with a different history of Facebook activity.

pro-Park groups were selected and added to our sample of seed pages for their relevance to Facebook activism groups in isolation from institutions. This purposive sampling was conducted by one of the authors, who discarded 12 pages for no apparent social/political cause found in group slogans and content, 25 pages for no post uploaded during the anti-Park movement period of interest, 33 pages for not being publicly available, 52 pages for being run by or affiliated with political institutions or established organizations, and 60 pages on which the posts were "liked" fewer than five times. By doing so, we could capture a temporally representative sample of activism group pages created to engage Facebook users with the anti-Park movement and its countermovement.<sup>6</sup>



**Figure 1. The production rhythms of posts on Facebook activism pages (left) and user comments on the posts (right) over the period of activism.**

From the sampled pages, information on who commented, when they commented, and in what words was obtained using the Rfacebook package that provided access to the Facebook API (Barbera,

<sup>6</sup> Note that there were some differences in the distribution of user activity between anti- and pro-Park groups, which made our sample of Facebook pages biased toward the former group. In this anti-Park group of activism pages, the most dominant page was <http://www.facebook.com/togetherson>, which elicited 103,640 comments; however, there were only 13 pages (8%) with more than 1,000 comments, and most of the pages (124 pages) were small-scale pages with fewer than 100 comments. By comparison, the majority of the pro-Park group had more than 1,000 comments (nine of 13 pages in total): The top three pages were <http://www.facebook.com/PaggeunhyeDaetonglyeongsidaeleulWihayeo>, which elicited 10,181 comments; <http://www.facebook.com/peoplepower2017>, with 7,404 comments; and <http://www.facebook.com/koreaauick>, with 6,656 comments. This different pattern of group organizing between anti- and pro-Park activism networks was reflected in the sample. The full list of Facebook pages included in our sample is available on request to the authors.



Piccirilli, Geisler, & van Attevelde, 2017); in doing so, all posts published during the time period and their associated comments were retrieved and archived from the selected pages. Figure 1 shows longitudinal changes in daily activities (posting and commenting) on both activism pages during the time period of the social movements: the anti-Park activism increasingly intensified pressure on the impeachment vote of the National Assembly on December 3, 2016, and subsided gradually afterward; by contrast, the pro-Park activism revved up to overturn the National Assembly's decision at the Constitutional Court on March 10, 2017. The different trends of Facebook engagement should be interpreted in relation to the contextual understanding of the movements.

### ***Computational Text Analysis***

Our computational text analysis drew on centering resonance analysis (CRA), which represents the content of large sets of texts by analyzing a complex discursive network of word co-occurrences, given the extent to which influential words bridge between other words in a single utterance (Corman, Kuhn, McPhee, & Dooley, 2002; Papacharissi & de Fatima Oliveira, 2012). This computer-assisted text analysis "describes the extent to which words are prominent in creating a structural pattern of coherence in a text" (Corman et al., 2002, p. 179) by locating the "in-between" position of the words in a co-occurrence network. In particular, this automated content analysis seeks to find nouns (or noun phrases) that link conceptual clusters together and thus help organize the whole; thus, it allows for surveying rich and complex structures in word networks to the extent that the influence of certain words in utterances is dependent not just on their frequency but also on their location in the semantic network structure. That is to say, this representational technique is structurally sensitive to a semantic network because "it accounts for all likely chains of association among the words that make texts and conversations coherent" (Corman et al., 2002, p. 181). Therefore, for contrasting the connective logic of meaning making around anti-Park activism with the semantically coherent structure of collective discourse in the pro-Park movement, our CRA technique is a more suitable method for weighing network structures than alternative text analysis approaches, such as a keyness analysis, which examines which words are given priority in a corpus.

Using the CRA technique, the comments on Facebook activism were analyzed as follows. The unit of analysis was a comment, not a user ID of the comment, to consider the agency of each utterance in the formation of activism discourse on Facebook. Next, anti- and pro-Park corpora of the comments were preprocessed and tokenized through the KoNLP package (Jeon & Kim, 2016). For Korean natural language processing, in particular, we used Hannanum as a morphological analyzer, and this part-of-speech (POS) tagger allowed us to extract nouns (including pronouns and numerals) used to name objects in comments on posts in relation to the given issue. This preprocessing method rests on the idea that the semantics of parts of discourse in Korean is effectively captured through nouns with lexical meaning (Liu & Wang, 2017). Especially, most nominal forms of Korean words turn into verbs (or predicates) when they conjoin with particles/postpositions; therefore, noun extraction is "a highly effective method for distilling documents to words with high information load" (C. Lee, 2019, p. 86). A detailed summary of the tokenized data is presented in Table 1.

Using the extracted nouns, we constructed each semantic network from anti-Park and pro-Park activism to examine how discursive engagement in the movement was organized during the affectively

charged moments. For the semantic network, a series of word co-occurrences were therefore accumulated (Corman et al., 2002), to compute how many times each pair of words appeared in the same comment together: by “co-occurrence,” we refer to the incidence of any pair of words appearing together on the same comment, and its strength was weighted by counting co-occurrences. And the moments of affectively charged activism were captured by focusing our analysis on the top 10 days of the most dynamic Facebook engagement.<sup>7</sup>

**Table 1. Summary of Tokenized Words With POS Tags.**

	Anti-Park activism	Pro-Park activism
No. of total ejels tokenized	2,189,228	715,611
No. of nouns	762,965 (34.85%)	267,599 (37.39%)
No. of predicates	429,546 (19.62%)	124,004 (17.33%)
No. of modifiers	177,806 (8.12%)	47,185 (6.59%)
No. of interjections	18,668 (0.85%)	4,206 (0.59%)
No. of particles/postpositions	231,539 (10.58%)	87,120 (12.17%)
No. of endings	571,520 (26.11%)	167,910 (23.46%)
No. of affixes	145,048 (6.63%)	53,999 (7.55%)
No. of pages	159	13
No. of comments	144,639	29,133

*Note.* *Ejels* refer to morphological units peculiar to the Korean language, which are classified into (1) nouns (common nouns, proper nouns, bound nouns, pronouns, and numerals); (2) predicates (verbs, adjectives, and auxiliary predicates); (3) modifiers (determiners and adverbs); (4) interjections; (5) particles/postpositions; (6) endings; and (7) affixes (prefixes and suffixes).

Subsequently, we compared the anti- and pro-Park semantic networks to learn how influential the words were in the network by bridging the shortest path between two other words that were not used together directly in the same comment (“betweenness”). This centrality metric for social-network analysis is useful for extracting linguistic meanings from semantic networks, insofar as betweenness as a method for CRA indicates the brokerage potential of words or noun phrases for bridging and coordinating different subjects or objects of utterances (Corman et al., 2002). A corpus of Facebook comments was thus folded into a network of interconnected words, using a Sankey diagram in the networkD3 package, to explore the context of semantic structures.<sup>8</sup> The visualization technique was useful for showing linkage weight between the words by positioning the words with high betweenness centrality at intermediate stages.

<sup>7</sup> The selected top 10 days of Facebook engagement were November 5, 19, 26, 28, 29, and 30, 2016, and December 1, 3, 4, and 9, 2016, for anti-Park activism. For pro-Park activism, the top 10 days were December 21, 22, and 24, 2016, February 9, 22, and 27, 2017, and March 4, 5, 8, and 10, 2017.

<sup>8</sup> Although Sankey diagrams are frequently used for a directed graph consisting of proportional flows from sources to targets, the method is also useful for creating an undirected graph to show which nodes are prominent in their interconnections with each other.

To compare the two semantic networks, furthermore, we randomly selected 1,000 comments from the top 10 days of engagement in each activism group. Their underlying structures of complex linguistic systems were then analyzed, grouping words together in community clusters based on the relationships between them. Doing so allowed us to examine whether the co-occurrence patterns of the words were generalizable to a broader period of social movement. In particular, the linguistic community structure of each semantic network was explicated through the community-detection algorithm for modularity optimization proposed by Blondel, Guillaume, Lambiotte, and Lefebvre (2008). This optimization task calculates the maximal value of modularity over possible divisions of a network into communities (Newman, 2006). "Communities" here refers to linguistic clusters of words that frequently co-occur in the same comment while rarely co-occurring with the words in other communities. Therefore, the community-detection algorithm was useful for assessing the extent to which the semantic network was organized around diverse repertoires of word relations. Especially, insofar as Gephi's community-detection algorithm compartmentalizes the global network into subcommunities, the words within the same community are linked to each other via strong ties and also linked to the words in other communities via weak ties.

Gephi was used for our visualization of the community structure to give a readable shape to the network's complex structure; in particular, the OpenOrd layout algorithm was employed to create a force-directed movement that distinguishes clusters among a large number of nodes. The algorithm operates based on Fruchterman-Reingold's algorithm, with 2,000 iterations for each network visualization (Martin, Brown, Klavans, & Boyack, 2011).<sup>9</sup> By doing so, the affectively rendered structure of the co-occurrence network was examined with regard to the function of words that bridge pluralistic clusters of activism discourse.

## Results

Our data analysis unfolds in two parts to address the research questions. First, we carefully captured affectively charged moments of each activism by analyzing the comments made during the top 10 days for Facebook engagement. Specifically, we explored the affectively rendered structure of semantic networks in the discourse about anti- and pro-Park activism by constructing the co-occurrence network of words from the comments. By doing so, we found that the network structure manifested very different patterns of word relations (see Table 2).

---

<sup>9</sup> We thank an anonymous reviewer for the suggestion that our visualization of semantic networks with different community structures can be improved using the OpenOrd layout algorithm.

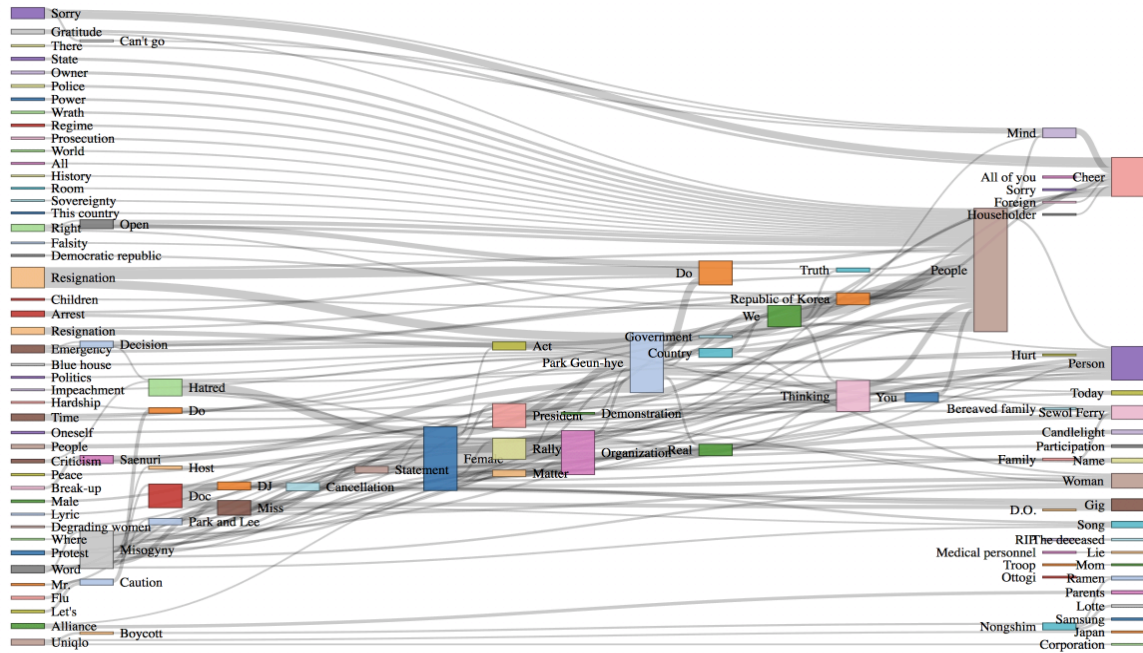
**Table 2. Semantic Network Structure of 1,000 Comments During Top 10 Activism Days.**

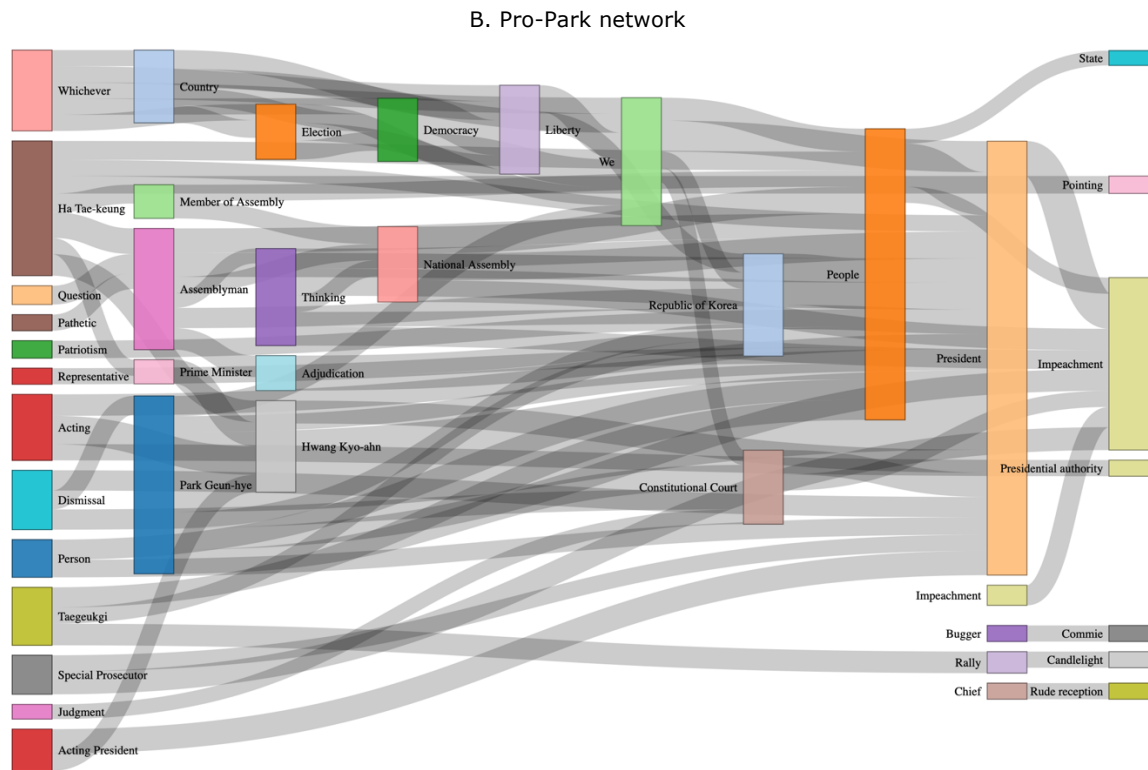
	Anti-Park network	Pro-Park network
<i>Network structure overview</i>		
No. of comments	1,000	1,000
Nodes (nouns)	2,199	4,028
Edges (co-occurrences)	13,366	95,675
Density	0.006	0.012
Transitivity	0.32	0.51
<i>Community structure</i>		
Connected components	113	55
Modularity	0.58	0.54
No. of communities	139	80

Note. Co-occurrence networks of words were derived from randomly sampled 1,000 comments that were made during the top 10 days of Facebook engagement in each activism group.

In particular, the data show that comments on the pro-Park pages contain many more nouns and their denser linkages than those on anti-Park pages. The pro-Park semantic network was roughly twice as dense as the anti-Park network (0.012 to 0.006), given that density measures the frequency of realized edges relative to possible edges. Also, insofar as transitivity captures the number of balanced triads to summarize the relative frequency with which connected triple nodes close to form triangles, the pro-Park semantic network is much more cohesive than the anti-Park network (0.51 to 0.32).

A. Anti-Park network





**Figure 2. Sankey diagrams for anti-Park and pro-Park semantic networks.**

Note. To visualize the prominent link weight of a word pair in a semantic network, only the edges above 10 weights were included. The thickness of each link indicates the number of connections between the two words. The height of each bar indicates the number of links with different words.

This contrasting pattern of word use is also observed in Sankey diagrams to visualize the weighted relationship between the words. The Sankey diagrams are to show the relative quantities of connections between nodes within a semantic network, given the weighted co-occurrence network in which a pair of words appears together in the same comment. When we limited our analysis to the edges weighted more than 10 to find the most prominent words and their connections in each semantic network, the opposing Facebook pages manifested different patterns of word use: Whereas the pro-Park network consists of 41 nodes and 85 edges, the anti-Park network consists of 112 nodes and 217 edges (see Figure 2).

More importantly, the anti-Park diagram rests on a decentralized structure with plural subcommunities, and there exist many important words that go the “in-between” position. As seen at the Graph A in Figure 2, there are some “in-between” words such as 여성 (female), 박근혜 (Park, Geun-hye), 조직 (organization), and 생각 (thinking) with a higher quantity of linkage than others, so that they extend across different clusters in the network. In addition, the prominent words invited many isolated words that are linked with no other words in the semantic network: The most connective word (highest degree) was 국민 (people), which has 838 links with 40 different words, followed by 박근혜 (Park, Geun-hye), with 675

links with 30 words, and 여성 (female), with 671 links with 27 words. This suggests that, through affectively rendered engagement in activism, Facebook comments coordinate multilayered, loosely connected personal frames to be crowdsourced in semantic organizing.

On the other hand, as shown at the Graph B in Figure 2, the pro-Park diagram shows a centralized structure around the word 대통령 (president), which has relatively infrequent links and with fewer words (271 links with 18 words). Moreover, the words with higher links tend to anchor in word connections rather than intermediating between other words. In this network, therefore, the linkage patterns of the words are more tightly interconnected with each other to form a coherent semantic structure.

Furthermore, we found that the two semantic networks differ strikingly in community structure. As shown in Table 2, on top of differing numbers of connected components (113 to 55), the anti- and pro-Park networks are organized around 139 and 80 clusters, respectively. Also, the composition of the communities is different. In the anti-Park network, the top 10 clusters account for about 69% of the words that are grouped in the community structure. But the pro-Park network has top 10 communities that explain the compartmentalization of about 79% of the words.<sup>10</sup> This community-detection algorithm allowed us to show how the opposing activism networks are organized around different structures of word relations.

The community structures of affective anti-Park and organizational pro-Park networks were visualized in Figure 3. Graph A shows the community structure of the affectively-rendered anti-Park network, in which diverse communities with distinct colors cross each other, so no prominent cluster emerges from the pattern of word co-occurrences. Rather, the pattern of word linkages creates a so-called random network in which any pair of nodes can be connected only through a handful of in-between nodes. This network structure indicates the manifestation of polysemous semantics in which the meaning of a word or concept can be defined by various words or concepts it connects to.

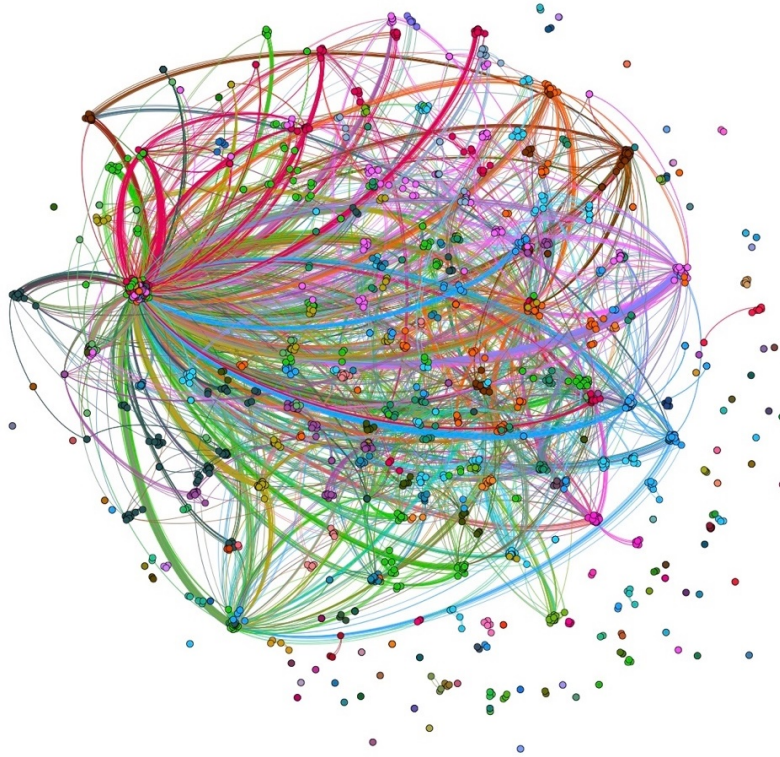
On the other hand, Graph B in Figure 3 shows some scale-free agglomerations of word relations in the organization-based pro-Park network of semantics. This community structure is characterized by a greater level of compartmentalization, in which nodes are tightly connected to each other in a cluster, but isolated from other clusters: This is far more noticeable than in the anti-Park network. Whereas there seems to be no distinct cluster of words tightly connected with each other in the anti-Park network, the pro-Park network maintains the segmented community structure in distinct clusters of word use. In this organization-based network, therefore, there exist many distinct clusters of word use that are distinct from each other. This structure of network representations for natural language semantics highlights the existence of some highly clustered neighborhoods that frame the meanings of words or concepts as coherent and established through collective practices.

---

<sup>10</sup> The top 10 communities detected in the word co-occurrence network of the anti-Park activism had the following proportions: 16.48%, 7.92%, 7.1%, 7.01%, 6.69%, 6.06%, 5.15%, 5.1%, 4.01%, and 3.51%. In the pro-Park activism, the top 10 communities were: 15.84%, 14.82%, 10.1%, 6.73%, 6.21%, 5.46%, 5.39%, 5.16%, 4.67%, and 4.54%.

Moreover, the method of modularity optimization in Gephi detected communities in networks by dividing the words of a semantic network into communities such that connections are dense within communities but sparser between them. This logic of community detection allowed us to differentiate between strong ties and weak ties in the network, insofar as the former indicates within-community edges and the latter between-communities edges. This means that the weak ties enable the semantic network to be organized heterogeneously around different ideas and frames of reference, whereas strong ties keep the network collectively homogeneous and coherent. Given the numbers of detected communities in the affect-driven anti-Park and organization-based pro-Park networks, it was expected that their semantic structures would display different roles of weak ties in the global graph.

#### A. Affect-driven anti-Park network



## B. Organization-based pro-Park network



**Figure 3. Community structures of anti-Park and pro-Park semantic networks during the top 10 days of Facebook engagement in activism.**

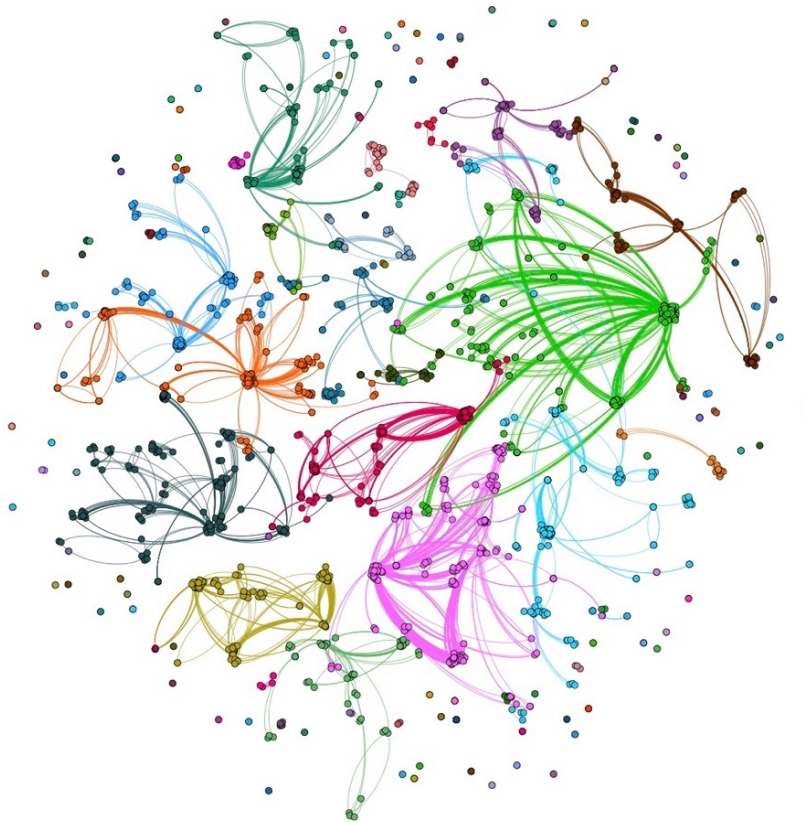
*Note.* Visualized is the co-occurrence network among words in the comments made during the top 10 days of Facebook engagement in activism. Dots indicate words, sized by their betweenness centrality scores. Nodes and edges are colored according to their community association.

The graphs in Figure 4 show that, when weak ties are removed from each semantic network, the two opposing networks have in common the existence of small worlds (clusters with strong ties). Of course, compared with the densely clustering structure of strong-tie word relations in the pro-Park network (Graph B in Figure 4), the network structure of anti-Park activism exhibits a contour of sparsely connected clusters that distinguish different boundaries from each other (Graph A in Figure 4). Nevertheless, it is certain that both activism networks maintain certain cohesive patterns of word use within strong-tie clusters on which the established meanings rest. Given the different network structures in anti-Park and pro-Park semantics, the shared structural characteristics of strong-tie networks suggest an important function of weak-tie word relations in the affect-driven activism that connect different clusters to each other. That is to say, the anti-Park network loosens a compartmentalization of semantics in such a way that weak-tie relations allow

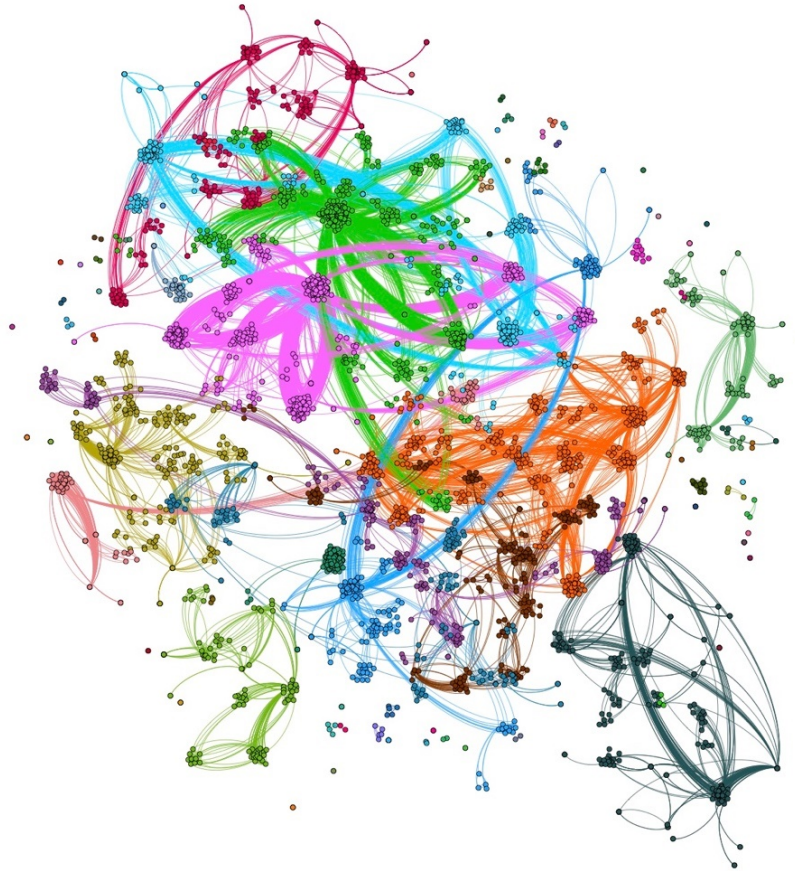


unregulated connections of diverse words. By contrast, it is clear that the organization-based activism manifests a limited function of weak-tie word relations insofar as the segmented clustering is observed even before strong ties are retained.

A. Affect-driven anti-Park network without weak ties



## B. Organization-based pro-Park network without weak ties



**Figure 4. Community structures of anti-Park and pro-Park semantic networks without weak ties during the top 10 days of Facebook engagement in activism.**

*Note.* Visualized is the co-occurrence network among words in the comments made during the top 10 days of Facebook engagement in activism. Dots indicate words, sized by their betweenness centrality scores. Nodes and edges are colored according to their community association.

Our comparison of community structures suggests that the organization-based pro-Park activism integrates Facebook engagement into a collectively coherent construction of word use through the primacy of strong ties. Although there are many more word relations (edges) in the pro-Park semantic network, their clustering is so coherent that the comments rest on densely interconnected ideas and frames of reference. On the contrary, the anti-Park semantic network facilitates the power of weak-tie relations in the construction of a heterogeneous community structure in which diverse ideas or concepts are frequently linked to each other not only within the same cluster but also across their clustering boundaries. Within this network of Facebook engagement, therefore, diverse semantic groups of word use are loosely connected to the discourse building of anti-Park activism. This difference corroborates the findings that (1) the organization-based pro-Park activism maintains collectively generated discourse in which word use is highly

tied to its cliché-like semantic context, but (2) the affect-driven anti-Park activism enables immediate but fleeting engagement in Facebook that bridges pluralistic, cross-domain expressions.

### Discussion

Affect is a liminal state of emotion that fills our time and space with potentials. Without repeating pre-given terms that still maintain "pre-authorized, pre-meant relations" between different groups of people, the heterogeneous language use on the anti-Park web page shows that the online space can be "immanent" enough to derive its "collective becoming" (Massumi, 2002, pp. 76–77). That is, the divergent patterns of word use on the anti-Park pages show that social media function as what Massumi (2002) calls an "immanent field of potential" (p. 78). According to Miall (1989), affect-driven language has three main characteristics: "1) self-referential, 2) cross-domain, and 3) anticipatory" (p. 61). In particular, the word "anticipatory" indicates that activism messages can invite the recipients' immediate preferences or sentiments toward a certain issue that keep their comprehension process ongoing, without having any clear cognitive content (Miall, 1989). Affective narratives, in this way, prestructure the meaning of the text. "Self-referentiality" implies that the meaning of the posted material is judged in relation to the user's own experiences (Miall, 1989). Affective news feeds on the anti-Park activism pages bring a myriad of self-referential responses in which people begin their own interpretative process based on prior experiences and feelings. Because this process is based on all the different perspectives of the commenters, activism discourse inevitably brings diverse access to private self-consciousnesses and personality variables (Fenigstein, Scheier, & Buss, 1975). Affect's "cross-domain" feature also echoes our finding that language is an essential enabler to blend fact, opinion, and sentiment to generate an intensified energy among social media users and infuse it with a live-streamed event (Papacharissi, 2015); this is the process in which affective discourses and participants' pluralistic expressions are collaboratively generated in online space.

In this study, we tested whether Facebook as a means of activism engagement characterizes a large-scale mobilization as affective, in that commenting on activism posts accommodates both personal and collaborative imaginations for individuals. As a result, we found that the anti-Park candlelight movement was affect driven, insofar as no clear organizational or ideological cleavages dominated its activism; rather, diverse communities of ideas were loosely connected with each other in semantic organizing through the bridging function of weak ties in the network. But its countering movement tended to rest on an organization-based collective action for which densely connected concepts/ideas maintained their isolation from each other in discourse building. We found this tendency by constructing semantic networks and removing weak-tie word relations between different communities from the complex graph. And this suggests that the community-detection algorithm can be a useful method for explicating the mechanism of political-identity formation on social media by analyzing unstructured text data.

The two groups of Facebook comments reveal different contours of semantic structures in their word co-occurrence networks. For the comments on the anti-Park activism pages, the structure of networks consists mainly of words that bridge different clusters of discourse, so that its system of word linkage is loosely connected but multilayered. This finding suggests that the affect-driven semantic network has a random-network structure within which personal action frames loosely opt into or opt out of activism engagement, so that diverse discourses are connected with each other. On the other hand, the comments

on the pro-Park pages exhibit a highly structured pattern of word relations in which their clusters are set apart from each other clearly. This organization-based network prioritizes strong-tie relations that strengthen collective frames of reference over weak-tie ones. That is to say, this network highlights a coherent pattern of discourse building in which semantically close words are frequently connected with each other but rarely conjoined with new ideas.

As shown in Figure 2, the two activism semantic networks show clearly different patterns of word relations. Whereas the anti-Park activism involves a diverse and heterogeneous structure of semantic clustering, the pro-Park activism is unified in its word use. For instance, the term *생각* (thinking) in the anti-Park activism comments connects with more diverse vocabularies that cut across diverse contexts of topics such as *국민* (people), *사람* (human), *여성* (female), *박근혜* (Park, Geun-Hye), *당신* (you), *단체* (group), *대통령* (president), *우리* (we), *자기* (self), *사람들* (people), *나라* (nation), *세월호* (Sewol ferry), *진짜* (real), *문제* (problem), *행동* (behavior), and *여혐* (misogyny). This word cluster mostly contains semantically unclear or broader word domains such as "real" and "group" or personal pronouns such as "we" and "you."

By contrast, the contour of the pro-Park semantic network shows a different trend of mobilization. Whereas the organization-based activism comments mostly intensify collective feelings of engagement, we found that dispersion of semantic meanings appears relatively slow, and they remain focused mostly on ideologically coherent terms, as listed in Figure 2. The word *생각* (thinking), in the comments on the pro-Park pages, is linked only to typically political terms such as *대통령* (president), *국민* (people), *탄핵* (impeachment), *국회* (National Assembly), *대한민국* (Republic of Korea), and *국회의원* (assemblyman). These words are all clearly indicative of the collective identity that the participants share and aim at.

### Conclusion

In this study, we demonstrated the importance of social media for connecting individual experiences and forming a public that enables networked activism. Especially, we targeted the anti- and pro-Park activism on Facebook to examine how the users' textual gestures on each site differed between movements. In doing so, we found that affect-driven human interactions produced more diverse and heterogeneous patterns of word use compared with organization-based collective action. Rather than constituting a homogeneous discursive process, affectively charged social movements occurred without their meaning being ideologically predetermined.

The power that an affective public attains online is of a liminal and transient nature, since it does not require a fixed or collectively shaped identity. A digitally mediated affective movement, however, generates game-changing effects that would not be available otherwise. Of course, the change can be slow and gradual. But social media play a dual-faceted function as "the sociality forms of a networked individualism" (Papacharissi, 2015, p. 101). And this function operates at both personal and collective levels, in this way rendering an ideologically open, semantically pluralized, and cross-referential interpretation of a social event. In other words, by integrating the users' personal and alive feelings with public ones, online languages allow users' performative storytelling to attain political relevance. However evanescent their voices might be in the current state, this study shows that the users' temporary voices, when integrated into digital networks, can be resilient enough to create large-scale social movements.

### References

- Austin, J. L. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132.
- Barbera, P., Piccirilli, M., Geisler, A., & van Atteveldt, W. (2017). Rfacebook: Access to Facebook API via R (Version 0.6.15) [Computer software]. Retrieved from <https://cran.r-project.org/package=Rfacebook>
- Benford, R. D., & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, *26*(1), 611–639.
- Bennett, W. L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication*, *58*(4), 707–731.
- Bennett, W. L., & Segerberg, A. (2013). *The logic of connective action: Digital media and the personalization of contentious politics*. New York, NY: Cambridge University Press.
- Bimber, B. A., Flanagin, A. J., & Stohl, C. (2005). Reconceptualizing collective action in the contemporary media environment. *Communication Theory*, *15*(4), 365–388.
- Bimber, B. A., Flanagin, A. J., & Stohl, C. (2012). *Collective action in organizations: Interaction and engagement in an era of technological change*. New York, NY: Cambridge University Press.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.
- Butler, J. (1997). *Excitable speech: A politics of the performative*. New York, NY: Routledge.
- Chen, H.-T., Ping, S., & Chen, G. (2015). Far from reach but near at hand: The role of social media for cross-national mobilization. *Computers in Human Behavior*, *53*, 443–451.
- Coretti, L., & Pica, D. (2015). The rise and fall of collective identity in networked movements: Communication protocols, Facebook, and the anti-Berlusconi protest. *Information, Communication & Society*, *18*(8), 951–967.
- Corman, S. R., Kuhn, T., McPhee, R. D., & Dooley, K. J. (2002). Studying complex discursive systems: Centering resonance analysis of communication. *Human Communication Research*, *28*(2), 157–206.

- Dalton, R. J., & Wattenberg, M. P. (2000). *Parties without partisans: Political change in advanced industrial democracies*. New York, NY: Oxford University Press.
- Diamond, L. (2010). Liberation technology. *Journal of Democracy*, 21(3), 69–83.
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43(4), 522–527.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297.
- Howard, P. N. (2015). *Pax technica: How the Internet of Things may set us free or lock us up*. New Haven, CT: Yale University Press.
- Howard, P. N., & Hussain, M. M. (2013). *Democracy's fourth wave? Digital media and the Arab Spring*. New York, NY: Oxford University Press.
- Jeon, H., & Kim, T. (2016). KoNLP: Korean NLP package (Version 0.80.1) [Computer software]. Retrieved from <https://cran.r-project.org/package=KoNLP>
- Jones, K. L., Noorbaloochi, S., Jost, J. T., Bonneau, R., Nagler, J., & Tucker, J. A. (2018). Liberal and conservative values: What we can learn from congressional tweets. *Political Psychology*, 39(2), 423–443.
- Lakoff, G. (2002). *Moral politics: How liberals and conservatives think* (2nd ed.). Chicago, IL: University of Chicago Press.
- Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2), 311–331.
- Lee, C. (2019). How are “immigrant workers” represented in Korean news reporting?—A text mining approach to critical discourse analysis. *Digital Scholarship in the Humanities*, 34(1), 82–99.
- Lee, F. L. F., & Chan, J. M. (2016). Digital media activities and mode of participation in a protest campaign: A study of the Umbrella Movement. *Information, Communication & Society*, 19(1), 4–22.
- Lee, S. H., & Fu, K. (2019). Internet use and protest politics in South Korea and Taiwan. *Journal of East Asian Studies*, 19(1), 89–109.
- Liu, W., & Wang, L. (2017). POS-tagging enhanced Korean text summarization. In D.-S. Huang, A. Hussain, K. Han, & M. M. Gromiha (Eds.), *Intelligent computing methodologies*:

- 13th international conference, ICIC 2017, Liverpool, UK, August 7-10, 2017, proceedings, part III* (pp. 425–435). Cham, Switzerland: Springer.
- Martin, S., Brown, W. M., Klavans, R., & Boyack, K. W. (2011). OpenOrd: An open-source toolbox for large graph layout. *Visualization and Data Analysis 2011*, 7868. doi:10.1117/12.871402
- Massumi, B. (2002). *Parables for the virtual: Movement, affect, sensation*. Durham, NC: Duke University Press.
- McAdam, D., Tarrow, S., & Tilly, C. (2009). Comparative perspectives on contentious politics. In M. I. Lichbach & A. S. Zuckerman (Eds.), *Comparative politics: Rationality, culture, and structure* (pp. 260–290). New York, NY: Cambridge University Press.
- Miall, D. S. (1989). Beyond the schema given: Affective comprehension of literary narratives. *Cognition & Emotion*, 3(1), 55–78.
- Morozov, E. (2011). *The net delusion: The dark side of Internet freedom*. New York, NY: PublicAffairs.
- Neiman, J. L., Gonzalez, F. J., Wilkinson, K., Smith, K. B., & Hibbing, J. R. (2016). Speaking different languages or reading from the same script? Word usage of Democratic and Republican politicians. *Political Communication*, 33(2), 212–240.
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), 8577–8582.
- Nielsen KoreanClick. (2016, October 19). *News consumption behavior change due to diversification of news platform*. Retrieved from [http://www.koreanclick.com/insights/newsletter\\_view.html?code=topic&id=417&page=1](http://www.koreanclick.com/insights/newsletter_view.html?code=topic&id=417&page=1)
- Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. New York, NY: Oxford University Press.
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information, Communication & Society*, 19(3), 307–324.
- Papacharissi, Z., & de Fatima Oliveira, M. (2012). Affective news and networked publics: The rhythms of news storytelling on #Egypt. *Journal of Communication*, 62(2), 266–282.
- Polletta, F., & Jasper, J. M. (2001). Collective identity and social movements. *Annual Review of Sociology*, 27(1), 283–305.
- Schneider, S. M., & Foot, K. A. (2004). The Web as an object of study. *New Media & Society*, 6(1), 114–122.

Williams, R. (1977). *Marxism and literature*. New York: Oxford University Press.

Woolley, J. K., Limperos, A. M., & Oliver, M. B. (2010). The 2008 presidential election, 2.0: A content analysis of user-generated political Facebook groups. *Mass Communication and Society, 13*(5), 631–652.