

## Discourses of Distance Education and COVID-19 in South Korean News between 2019 and 2021: A Topic Modeling Analysis

Kyungmee Lee, Lancaster University, UK, <https://orcid.org/0000-0002-9580-9026>  
Tae-Jong Kim, Korea Institute of Science and Technology, South Korea, <https://orcid.org/0000-0001-8087-2092>  
Aras Bozkurt, Anadolu University, Turkey, <https://orcid.org/0000-0002-4520-642X>  
Olaf Zawacki-Richter, University of Oldenburg, Germany, <https://orcid.org/0000-0003-1482-8303>  
Berrin Cefa Sari, University of Oldenburg, Germany, <https://orcid.org/0000-0002-8306-8574>  
Victoria I. Marín, University of Lleida, Spain, <https://orcid.org/0000-0002-4673-6190>

**Abstract:** This paper aims to identify dominant discourses of distance education that emerged in Korean society before, during, and after the COVID-19 outbreak in January 2019. To achieve the aim, the authors have conducted a Topic Modeling analysis of 8,865 news articles published by 54 South Korean media outlets between 2019 and 2021. As a result, five key topics and the top 10 keywords associated with each topic have been identified for 2019, 2020, and 2021, respectively—15 topics and 150 keywords in total for the three years. There have been meaningful shifts not only in the quantity of news articles published each year but also in the quality of conversations presented in those articles on distance education. This article carefully analysed and reflected both continuity and discontinuity of distance education discourses in Korean society. Based on the results, we have drawn four discussion points: a) the normality of DE discourses, b) the historicity of DE discourses, c) the maturity of DE discourses, and d) the partiality of DE discourses. Despite the specific cultural context from which the four points were drawn, the discussions offer valuable insights applicable to different cultural contexts.

**Keywords:** discourses, distance education, COVID-19, South Korea, Topic Modeling analysis, news Big Data analysis

### Introduction

Within a long-established field of distance education (DE), there have been popular discourses about the accessible (or flexible) nature of DE provision (Lee, 2021). The "anyone, anytime, anywhere" rhetoric effectively demonstrates the positive perception of DE (particularly online DE) prevailing in society. The number of online courses and students, especially in higher education and adult education contexts, has increased in the past decade. It has frequently been referred to as evidence of the increased accessibility of today's educational systems (Lee, 2017). Contrarily, there has also been a dismissive view of DE as inferior to face-to-face education, which is often considered a gold standard for all educational provisions. While such competing discourses of DE existed unresolved, the global outbreak of the COVID-19 has forced educational providers of all levels to move online—the rapid adoption of DE was not a matter of choice but survival in most cases (Bozkurt & Sharma, 2020, Lee et al., 2021).

Subsequently, an unprecedented level of public conversations and interest in different aspects of DE emerged during the COVID-19 pandemic. Such rapid and unprepared adoption of DE has also revealed both premises and challenges associated with DE provisions. At this moment, when more conversations on the post-pandemic education (including DE or blended learning practices) spring, it could be a meaningful attempt to revisit some of the changes in general perceptions of DE that occurred during the pandemic to envision a better and more realistic future. Thus, we have collected news articles about DE produced and distributed between 2019 and 2021—the period enabling us to trace the meaningful ruptures in dominant discourses about DE. We have conducted a Topic Modeling analysis of news Big Data published in the South Korean context to trace what has (not) been stated and discussed in those media texts before, during, and after the COVID-19 outbreak.

### Research Methods

#### Data Collection

This study utilised BIGKinds (<https://www.bigkinds.or.kr>) to collect news articles from 24 media outlets in South Korea. BIGKinds is a Big Data research solution offered by the Korea Press Foundation that contains about 70 million news contents published by 54 different media outlets collected from 1990 and onward. We first selected nine major newspapers with nationwide readership and nine other regional newspapers with the largest readership in each region to adjust for geographical differences. We also included two newspapers specialising in business news and four TV news outlets (see Table 1).

The Tenth Pan-Commonwealth Forum on Open Learning (PCF10)  
September 14-16, 2022, Calgary, Canada.



Table 1. Media outlets chosen for data collection

Type	No	Selected media outlets
National Newspapers	9	Kyunghyang Shinmun, Dong-A Ilbo, Munhwa Ilbo, Seoul Shinmun, Segye Ilbo, JoongAng Ilbo, Chosun Ilbo, Hankyoreh, Hankook Ilbo
Regional Newspapers	9	Joongboo Ilbo, Kangwondomin Ilbo, Jungbu Maeil, Chungcheong Today, Maeil Shinmun, Busan Ilbo, Jeonbukdomin Ilbo, Jeonnam Ilbo, Jemin Ilbo
Business Newspapers	2	Maeil Business, Korea Economic Daily
TV News Outlets	4	KBS, MBC, SBS, YTN

To establish a comparative perspective, we collected news articles published before and after the COVID-19 outbreak divided into three separate annual periods—2019, 2020, and 2021. We searched all news articles, including the word "distance education" or equivalent search terms in their titles, as follows:

- distance learning, distance teaching, distance course, online education, online learning, online teaching, online course, cyber education, cyber learning, cyber teaching, cyber course, e-learning, remote education, remote learning, remote teaching, remote course, untact education, untact learning, untact teaching, untact course

The initial search resulted in a total of 10,157 news. After removing duplicates, personal affairs and event announcements, we first reviewed 9,600 news titles. Additional 735 articles that focus on foreign news or the following content were excluded:

- specific stock items, corporate products, advertisements, and promotions
- individual celebrities' and companies' announcements and events
- student recruitments and job advertisements
- representations of celebrities' posts on social networking sites
- issues and accidents unrelated to distance education (e.g., school lunch)

The selection process resulted in 8,865 news articles—239 for 2019, 6,504 for 2020, and 2,122 for 2021, with a noticeable increase in the number of articles published in 2020 compared to ones in 2019. The first and second authors of the present articles (who speak Korean as their first language) were responsible for the initial search and selection process, and all results were cross-checked between the two.

### Data Analysis

We performed a Topic Modeling analysis to identify key topics in the collection of news articles and their changes across the three periods. Based on the assumption that the composition of topics is a latent variable to be inferred, Topic Modeling analysis uses a set of observed variables to infer the desired latent variables. Using a social network analysis software (NetMiner, <http://www.netminer.com>), we first filtered out all the words in the selected news texts, not in the noun form, via the programme's morphological analysis function. We also removed single-word expressions whose meaning could not be reliably defined without further context. A series of pre-processing measures were undertaken to make the analysis more reliable. For example, all the words with the Term Frequency-Inverse Document Frequency (TF-IDF) value of 0.1 or less were removed— stop words. TF-IDF value for each word is calculated by retrieving the frequency of the term in each document and the scaled inverse fraction of the frequency of documents containing the term. A high TF-IDF value indicates a strong relationship between the term and the associated documents, whereas a low value denotes commonly used expressions (e.g., distance, education, COVID, spread) that are less useful in distinguishing one document from another.

We first extracted all keywords frequently included in the collected news articles and reviewed from the 1<sup>st</sup> keyword "class" mentioned 54,017 times across the 8,865 news articles to the 500<sup>th</sup> keyword "group" mentioned 474 times. We created three keyword lists based on the review results (i.e., synonyms, exclusions, definitions) to refine our Topic Modelling analysis method. We used the Latent Dirichlet Allocation (LDA) algorithm ( $\alpha=0.1$ ,  $\beta=0.01$ ), which is a statistical text processing method that identifies the proportion of topics in the collected texts and the probability of each word belonging to any given topic. Re-sampling frequencies were set as 1000 in the process of identifying the latent topics, which resulted in five as the optimal number of topics to be discovered from each of the three periodic datasets. To evaluate the clustering performance of the analysis, we calculated the Silhouette coefficients, which were 0.910, 0.924, and 0.882 for the three periods, respectively (i.e., the silhouette coefficient close to 1 indicates effectiveness), and the coefficients decreased gradually as we increased topic numbers. We repeated the Topic Modeling analysis and reviewed the results (i.e., five identified

topics and 50 top keywords most probably associated with each topic) three times to further refine our analysis by updating the keyword lists. We have identified and named the top five topics from each period and extracted the top ten words with the highest weights for each topic.

## Results

### 2019. Discourses before the COVID-19 outbreak

The first topic identified in 239 news articles concerning distance education (DE) published in 2019 is the adoption of DE for civil defence training. In South Korea, being exposed to the political tension and threat posed by the North Korean government, there has been an ongoing effort to increase national civil defence capabilities, with an annual budget of 43 trillion won (\$40 billion) allocated in 2018. Some of the regional civil defence training programmes that had long been operated using physical facilities were moved online in 2019. The top 10 keywords associated with the topic (Table 2) suggest that it was aimed to increase the "cost-effectiveness of such programmes and relieve the "burden" of both training organisers and participants. Subsequently, the ID "authentication" of participants in the online "assemblies" appeared to be one of the major interests discussed in the related news.

Topics 2 to 5 each were concerned with DE provisions at different levels of education as Topic 2 about lifelong learning, Topic 3 school learning, Topic 4 advanced skill training, and Topic 5 higher education. Based on the extracted keywords and associated news articles on each topic, we could see the meaningful differences in the public discourses concerning DE provisions in different educational contexts. For example, in lifelong learning contexts, participating in DE often provides "internationally" recognised "certificates" (a "psychological" "counsellor" certificate being the most frequently mentioned in the news articles). Along with the increased number of open courses, there were conversations about some of the DE programmes being "free of charge", while those programmes experienced security problems, including the "leakage" of the "personal: information of their "members".

When it comes to DE in school settings, the "Ministry of Education" and the regional/local "Offices of Education" appeared to be leading parties in the adoption and implementation of DE. "Students" and "teachers" also appeared as the main subjects in DE, and the need for teachers' "professional training" was much discussed. In 2019, "English" was the main "subject" utilising DE, which was perceived as "innovative" and exceptional practice. In the context of DE for the advanced skill training for the "safety", "health", "medical", and "social welfare" related subjects, more advanced technology such as small "drone" was mentioned. Managing and reducing associated "risk" to practical skill training at workplaces like "hospitals" appeared to be one of the core concerns.

There are 19 "cyber universities" specialising in DE provisions across Korea. In 2019, most news articles concerning DE in higher education featured the achievements of those universities. As keywords suggest, setting up "international" "collaboration" "agreement" with "foreign" organisations like universities was reported as top news. "Investments" in the new online learning platforms (often "start-up" EdTech companies) were considered important for "business management". Thus, the for-profit and special nature of DE was salient in the higher education context.

### 2020. Discourses during the COVID-19 outbreak

Unsurprisingly, with the outbreak of COVID-19 in January 2020, the number of news articles concerning DE provisions increased dramatically (from 239 in 2019 to 6,504 in 2020). DE quickly gained its normality, legitimacy, and necessity across all educational sections in Korea. The five topics identified in the 2020 news articles can be found in Table 3. Each was named as follows: university tuition return movement, disadvantaged population, business and innovation, test-taker and kindergartener, and challenges in primary school.

The most frequently recorded issue was university students' dissatisfaction with their universities' poor quality of DE, especially during the Spring 2020 "semester". Since most campus-based universities (unlike cyber universities) did not have previous experiences and required skills and infrastructures for DE, the perceived quality of DE provisions during the first semester after the COVID-19 outbreak was particularly low. In fact, most universities postponed the "start of the term" from March to April, "extending" the Winter break for a few weeks hoping that the spread of the virus would cease soon. Subsequently, there was student protest and movement for "tuition fees" "return". Especially students in specific disciplines that require much "practical training" and hands-on learning voiced out more. It is worthwhile to mention that many universities refunded at least a small portion of the tuition fees in the first semester.

The second most frequently discussed topic in 2020 concerns the challenges experienced by disadvantaged populations while engaging with DE. Diverse groups of "disadvantaged" "children" and struggling "adolescents" were featured in news articles, including students from "multicultural families" (in Korean settings, often referred to the families one or both parents are recent immigrants, often without financial, social, and cultural capitals), students with "disabilities" and students from low "social class". Given that school lunches and dinners are entirely free of charge in all public schools in Korea moving classes online caused significant damage to these groups of students—not only academically and physically. Many news articles argued for effective "welfare" measures and "social support" for such families.

Ironically, the tone of news articles associated with Topic 3, "Business and innovation", was rather positive and optimistic despite being published at the same time as Topics 1 and 2 above. Many articles discussed the opportunities the COVID-19 pandemic had provided in terms of the rapid development of the "EdTech" "business" and "market" and EdTech-driven "innovation" in education. It was proclaimed that the new "Digital Era" (and the "future" of education) had opened. Specific "companies" leading the changes, collaborating with government organisations and educational institutions, and providing new "solutions" were mentioned in those articles.

Topic 4 concerns the specific impacts of DE on the experiences of college entrance test-takers and pre-school kids. The implications of the continuous shifts in "social distancing" measures and policies on these two learner groups were discussed and reported as frequently (and soon) as national emergency stages changed. As the college entrance test is notoriously high-risk in Korean society, in which about 80% of high school graduates go to universities, issues related to the educational concerns created by the closure of "private institutes" and how to organise "test sites" that were usually "densely populated" were frequently discussed. In addition, the closure of pre-schools, which are common and affordable childcare providers in Korea, caused a range of social issues, especially for working parents and low-income families. Many news articles pointed out that infants (unlike other student groups) were unlikely to benefit from DE provisions.

An extensive set of news articles also reported diverse challenges observed and experienced in primary school contexts. The two most frequently mentioned issues were the "Internet connection" and the "digital divide". Following the Ministry of Education's guidance, all primary schools in Korea adopted a "two-way" communication (or synchronous) model for their DE provisions, using telecommunication tools such as Zoom and Google classroom during the pandemic. In this context, a lack of Internet connection (absence of "devices") at home appeared to be a significant social and educational problem—subsequently, issues of low attendance rates among students from disadvantaged families. Teacher professional training and student difficulties with concentration were also mentioned.

### **2021. Discourses after the COVID-19 outbreak**

As the COVID-19 pandemic continued, similar issues and concerns were still discussed (unsolved) in the news articles on DE in 2021. Nevertheless, there were new topics emerged as well. As Table 4 demonstrates, "Vaccine and school" were the most frequently discussed topic in 2022. The new availability of "vaccines" enabled different educational institutions, including "private institutes", "middle schools" (high schools), and "pre-schools", to resume in-person classes and childcare services in a more sustainable and consistent manner. Thus, the changes in the social distancing "measures" were continuously updated and reported in the news articles. The topic of the challenges experienced by disadvantaged populations while engaging with DE remained the second most frequently discussed topic in 2021. However, one noticeable shift from the discourse of 2020 is the appearance of such keywords as "Budget" and "Internet" in the same discourse of 2021. This suggests that there had been some improvement in the welfare policies for these groups, evident in the specific budget conversations and the strategies to provide "device" and "Internet"—in most cases, free of charge.

Another new topic previously not mentioned much was the side-effects of DE, namely. A large number of social and criminal incidents like "cyberbullying" among students on online platforms (including social networking sites) and "invasion" of "teacher authority", including verbal and sexual insults towards teachers during online classes, were reported. "Personal" "information" leakages and subsequent adoption of information "protection policies" were also mentioned. Unlike 2020, in 2021, there were more long-term and mid-term damages created by the prolonged period of DE provisions.

The topic of "Business and innovation" remained relatively unchanged except for the appearance of new keywords "smart" and "BigData", which suggests the maturity of the discourses. While such business-oriented

discussions in 2020 were still reactive to the COVID-19 outbreak, the same discussions in 2021 sounded more progressive and proactive—that is, more future-oriented. The topic of the college entrance test also remained one of the five topics in 2021. However, as both the COVID-19 vaccine and test were made available in 2021, the conversations were more specifically concerned with the measure for COVID-19 patients who had to self-isolate on the test day. Subsequently, how to organise test sites at hospitals and how to conduct test-takers' identification were discussed.

### Discussion and Conclusion

We have conducted a Topic Modeling Analysis of 8,865 news articles published by 54 South Korean media outlets between 2019 and 2021 to identify the major discourses on DE that emerged before, during, and after the COVID-19 outbreak in January 2020. As a result, five key topics and the top 10 keywords associated with each topic have been identified for 2019, 2020, and 2021, respectively. Based on the results, we have drawn four discussion points.

Firstly, as previously discussed in the recent DE literature, DE had long been a marginal and abnormal educational delivery method as in-person educational delivery had been perceived as normal (and often more effective) in most learning contexts. Of course, there was a noticeable exception for lifelong learning provisions where the beneficiaries were usually adults with other professional and social responsibilities and potentially preferred the accessible and flexible DE to in-person education. The results from 2019 news articles effectively demonstrated this long-standing trend as DE was considered useful for limited and specialised educational contexts, including civil defence training, lifelong learning, and advanced skill training. DE in schools and higher education were still considered something different and unique—either the innovative or commercial nature was stressed in the selected news article. Not surprisingly, in 2020 and 2021, as DE became the primary (and often the only) educational delivery method, **the normality of DE discourses** has quickly been established in the discourses of all educational levels.

Secondly, **the historicity of DE discourses** is clearly demonstrated by the results. The unique characteristics of existing educational systems and the specific conditions of cultural and social background in Korean society have shaped the development of both DE practice and discourses. Evidently, much attention was paid to the adoption of DE for civil defence training, which is one of the major areas of educational provisions that may not be observed in many other countries. Also, during the COVID-19 pandemic, there has been an ongoing discussion about the college entrance test and test-takers that show the importance of the high-stake exam in Korean society. Therefore, although the DE as a new normal can be seen as a global phenomenon during the pandemic, it is important to note how DE has been adopted and discussed is historically shaped and closely influenced by the social and educational background of each given society.

Thirdly, we could see **the maturity of DE discourses** in 2021—the second year of the pandemic. In 2019, the DE conversations were rather superficial and simplistic as many news articles published in 2019 limitedly reported positive aspects (or potentials) of DE, such as being free of charge, accessible, and innovative. Business-oriented and for-profit nature of DE provisions by specific educational institutions such as cyber universities was also much celebrated. Overall, there was a lack of criticality and balanced perspective encompassing both premises and challenges associated with DE practices. The immaturity of DE discourses before the pandemic was more likely an outcome of a lack of social and public interest in DE practice instead of intentional propaganda of specific parties. DE discourses in 2021 compared to 2019 (even compared to 2020 when much attention went into maintaining educational systems by adopting DE) were much more sophisticated. More in-depth discussions on the challenges experienced by disadvantaged student groups and the welfare and support strategies (with the required budget) appeared in 2021. Also, the side-effects and negative consequences of adopting DE for a prolonged period were carefully unpacked. That is, long-term and mid-term damages potentially caused by a long absence of in-person teaching in different educational settings were identified and discussed in 2021.

Finally, **the partiality of DE discourses** was also noticed. In both 2020 and 2021, throughout the COVID-19 pandemic, there was a clear divide between the DE discourses concerning the subjects of disadvantaged learners and the subjects of business-oriented actors. The nature and tone of each side of the discourse were dramatically distinctive, demonstrating the unfortunate inequality persistent in our society. Simply put, we can see who has been suffering from the adoption of DE following the COVID-19 outbreak and who has been seizing this lifetime opportunity to produce a fortune. While it is not important to make a list of those disadvantaged and

advantaged being separated on the other side of each other, we would argue that the critical awareness of such division and continuous efforts to close the gap (in particular, to address the damages of the disadvantaged) should remain ever after the pandemic ends.

## References

- Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian Journal of Distance Education*, 15(1), 1-6. <https://doi.org/10.5281/zenodo.3778083>
- Lee, K. (2021). Openness and innovation in online higher education: A historical review of the two discourses. *Open Learning*, 36(2), 112-132. <https://doi.org/10.1080/02680513.2020.1713737>
- Lee, K. (2017). Rethinking the accessibility of online higher education: A historical review. *The Internet and Higher Education*, 33, 15-23. <https://doi.org/10.1016/j.iheduc.2017.01.001>
- Lee, K., Fanguy, M., Bligh, B., & Lu, S. (2021). Adoption of online teaching during the COVID-19 Pandemic: A systematic analysis of changes in university teaching activity. *Educational Review*. <https://doi.org/10.1080/00131911.2021.1978401>

Table 2. Extracted topics and keywords in 2019

Category	Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
Topic	DE for civil defence training		DE for lifelong learning		DE in school		DE for skill training		DE in higher education	
1	Civil defence	0.066	Certificate	0.027	Student	0.047	Safety	0.026	University	0.033
2	Training	0.048	Free of charge	0.027	Class	0.035	Health	0.020	Agreement	0.015
2	Authentication	0.024	Psychology	0.022	School	0.032	Society	0.016	Cyber university	0.014
3	Emergency call	0.010	EdTech	0.020	Ministry of Education	0.014	Certificate	0.014	Foreign	0.011
4	Cost-effectiveness	0.009	Counsellor	0.019	Professional training	0.013	Welfare	0.012	Platform	0.010
5	Burden relief	0.009	Personal	0.016	Teacher	0.011	Drone	0.012	Start-up	0.008
6	Subject	0.008	Leakage	0.011	Office of Education	0.010	Medical	0.011	Business management	0.008
7	Safety	0.007	Member	0.011	English	0.009	Exam	0.010	Investment	0.008
8	Assembly	0.006	Employment	0.010	Subject	0.009	Hospital	0.007	Internationalisation	0.007
10	Place	0.006	International	0.007	Innovation	0.007	Risk	0.007	Collaboration	0.007

Table 3. Extracted topics and keywords in 2020

Category	Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
Topic	University tuition return		Disadvantaged population		Business and innovation		Test-taker and childcare		Challenges in primary school	
1	Professor	0.034	Child	0.013	Business	0.012	Private institute	0.025	Connection	0.012
2	Semester	0.019	Adolescent	0.012	Company	0.012	College entrance test	0.02	Two-way	0.009
2	Tuition fees	0.017	Disabled student	0.009	Future	0.01	Preschool	0.02	Teacher	0.009
3	Term start	0.015	Multicultural family	0.008	Digital	0.008	Test	0.013	Professional training	0.009
4	Extension	0.011	Welfare	0.008	Era	0.007	(social distancing) Stage	0.013	Device	0.008
5	Practical training	0.011	Social class	0.006	Innovation	0.006	Superintendent	0.011	Primary school	0.008
6	University	0.01	Disadvantaged	0.005	EdTech	0.006	Emergency	0.009	Divide	0.008
7	Subject	0.008	Special education	0.005	Collaborative	0.006	Densely populated	0.007	Attendance	0.008
8	Return	0.006	Support	0.005	Market	0.005	Test site	0.007	Internet	0.007
10	Exam	0.006	Family	0.005	Solution	0.004	Test-taker	0.006	Concentration	0.004

Table 4. Extracted topics and keywords in 2021

Category	Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
----------	---------	--	---------	--	---------	--	---------	--	---------	--

<i>Topic</i>	Vaccine and school		Disadvantaged population		Side-effects of DE		Business and innovation		College entrance test	
1	Private institute	0.014	Child	0.014	Cyber	0.023	Company	0.011	College entrance test	0.033
2	Preschool	0.012	Adolescent	0.011	Safety	0.018	Digital	0.01	Test site	0.02
2	Vacation	0.011	Welfare	0.008	Invasion	0.016	Information	0.008	Test	0.018
3	Vaccination	0.011	Device	0.007	Teacher authority	0.012	Future	0.008	Test-taker	0.015
4	Middle school	0.01	Disabled student	0.007	Bullying	0.008	Innovation	0.007	Test identification	0.011
5	Infection	0.009	Social class	0.007	Information	0.007	Industry	0.006	Test takes	0.008
6	Vaccine	0.008	Internet	0.006	Protection	0.006	Smart	0.006	Test preparation	0.008
7	School starts	0.007	Support	0.005	Policy	0.006	Collaboration	0.005	Hospital	0.006
8	Measure	0.006	Difficulty	0.005	Personal	0.005	Market	0.005	Isolation	0.006
10	In-person class	0.006	Budget	0.005	Damage	0.005	BigData	0.005	COVID test	0.005