Korean Journal of English Language and Linguistics, Vol 23, September 2023, pp. 696-712
DOI: 10.15738/kjell.23..202309.696

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Technology-Enhanced Suggestions for Customizing CLIL Materials: With the Topic of Virtual Humans*

Hyunyoung Moon· Eun-Young Kwon· Kyeongmin Woo (Korea Military Academy)

This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: July 6, 2023
Revised: July 24, 2023
Accepted: September 14, 2023

Hyunyoung Moon (1st author)
Assistant Professor, Dept. of English, Korea Military Academy
E-mail: hymoon2016@gmail.com

Eun-Young Kwon
(corresponding author)
Associate Professor, Dept. of English, Korea Military Academy
E-mail: eykwonkma0524@gmail.com

Kyeongmin Woo (co-author)
Assistant Professor, Dept. of English, Korea Military Academy
E-mail: martincan@kma.ac.kr

* This paper was supported by 2023 research fund of Hwarang-dae Research Institute.

ABSTRACT

Moon, Hyunyoung, Eun-Young Kwon and Kyeongmin Woo. 2023. Technology-enhanced suggestions for customizing CLIL materials: With the topic of virtual humans. *Korean Journal of English Language and Linguistics* 23, 696-712.

This study investigated the intricate dynamics surrounding the public reception of the virtual influencer Rozy in South Korea, utilizing a text-mining approach on data extracted from popular platforms such as Naver, YouTube, and Instagram. Drawing from a vast corpus of user comments, the research revealed significant differences in public sentiments and narratives, a phenomenon distinctly modulated by the individual characteristics of each platform and varying degrees of user anonymity: Naver, offering the highest level of user anonymity, resulted in a lot of critical comments towards the virtual influencer, while YouTube centered more on technological discourse and Instagram gathered positive engagements. Additionally, this study showcased the development of an English wordlist by utilizing Python libraries for collecting corpus data from platforms such as Google News, YouTube, and Quora, which can be useful sources for learners in CLIL classrooms, fostering deeper engagement with the emergent influencer culture. It emphasized the necessity of a nuanced approach in data collection and analysis, recognizing the distinct characteristics of each platform and the content they host, thereby contributing significantly to CLIL pedagogy through the creation of an English wordlist developed from authentic materials, fostering enhanced learner engagement.

KEYWORDS

CLIL, virtual influencer, Rozy, text mining analysis, Python libraries, authentic materials, Wordlist, discourse analysis

© 2023 KASELL All rights reserved

1. Introduction

With the advent of the Fourth Industrial Revolution, terms like ‘artificial intelligence (AI),’ ‘virtual humans,’ and ‘metaverse’ have become increasingly prevalent. The trend extends to academic settings; textbooks, especially those for college English courses, consistently feature topics related to technology, future technologies, and AI, underscoring the blend of current trends and student interests (Chase and Johannsen 2019).

However, a gap emerges when considering that many Korean students gravitate more towards domestic issues than foreign cases, yet most university English textbooks, being foreign publications, often overlook domestic situations. This emphasizes the importance of understanding complex concepts or societal phenomena within one's cultural context, preferably in one's native language. Content and Language Integrated Learning (CLIL) is a pedagogical approach that emphasizes both content and language. By actively utilizing scaffolding, this approach provides necessary language support through concepts such as translanguaging and code-switching (Tsou 2018).

The primary goal of this research is to examine how Koreans perceive a particular subject, with the intention of aiding CLIL educators and learners in Korea to activate their existing knowledge structures, and seek to help the learners express their views with regard to the topic in English. To achieve this, an analysis of social media big data was conducted to understand Koreans' perceptions of Rozy, a virtual human and influencer in South Korea, as a preliminary step in schema activation, and practical methods utilizing Python libraries for instructors to create vocabulary lists based on authentic textual materials from various platforms were proposed. Consequently, this research aims to offer pedagogical suggestions to domestic CLIL instructors who teach AI-related content.

2. Literature Review

2.1 Virtual Influencers and the Rise of Rozy in South Korea

The emergence of virtual influencers, a vibrant aspect of modern advertising and entertainment, can trace its roots back to animation characters of yesteryears. Initially, these characters found utility as tools for advertising in the 1940s, but the genesis of virtual characters performing as “virtual celebrities” began in the mid-1990s in Japan (Conti, Gathani and Tricomi 2022, p. 2). This period witnessed a transition of these entities from mere computer-generated imagery (CGI) characters to influential figures hosting offline events and festivals (Black 2008, Rahmi, Rahmat and Saleha 2018, Kobayashi and Taguchi 2019). However, it was only in 2016 with the introduction of Lil Miquela, a virtual entity with a human-like appearance and personality, that the concept of a “virtual influencer” truly took shape. Lil Miquela opened up a new era, actively participating in marketing and social campaigns, predominantly on Instagram, a platform most frequented for influencer marketing (Conti, Gathani and Tricomi 2022, Drenten and Brooks 2020, Rodrigo-Martin, Rodrigo-Martin and Munoz-Sastre 2021).

South Korea swiftly embraced this phenomenon, becoming a hotspot for virtual influencers. In August 2022, CNN highlighted the exponential growth of virtual influencer marketing in Korea, underscoring figures like Rozy, who amassed a following ranging from 32,000 to 3.9 million (Yeung and Bae 2022, Rasmussen 2022). Rozy, a brainchild of Sidus Studio X, made a groundbreaking debut in a commercial by ShinhanLife on July 1, 2021. This CGI entity, characterized by her unique features and engaging persona, significantly deviated from conventional Korean beauty standards, fostering a substantial following and an impressive number of views on YouTube.

Rozy, despite being an entrancing figure, was not the first virtual being in Korea, a title held by Adam who debuted in 1998. Yet, she marked a departure from the earlier virtual characters by establishing a vivid, interactive

© 2023 KASELL All rights reserved

presence on social media platforms, regularly posting content and fostering an almost human-like connection with her audience (Ahn and Lee 2021). Despite her realism causing unease amongst a faction who perceive her as a precursor to the replacement of human jobs (Hiort 2022), Rozy’s intrigue has encapsulated millions, becoming a sought-after entity for corporate collaborations and lifestyle promotions.

The advent of virtual figures like Rozy has stirred varied sentiments amongst the populace, raising questions about their reception—are they perceived as threats or fascinating innovations? Addressing these queries necessitates a comprehensive analysis of discourse on popular Korean social media platforms, utilizing text-mining methodologies to gauge the breadth and depth of public sentiment (Park 2022, Lee 2021, Park et al. 2022, Lee and Ma 2022). Existing studies, although predominantly focused on specific characteristics of virtual influencers like fashion and physical attractiveness, have begun venturing into analyses of real consumer responses. One notable study employed Python and Textom to evaluate reactions to Rozy on Instagram, reporting a majority of positive responses (Park and Koo 2022).

As the influence of virtual personalities continues to expand globally, with many like APOKI, Lechat, Rina, and others amassing significant popularity, comprehending the wider societal implications and responses becomes imperative. Rozy, in particular, serves as a focal point for evaluating the dynamic between virtual influencers and young audiences, especially considering her resonance amongst a specific group of young people and the consequent research dedicated to studying her (Lee and Ma 2022, Hwang and Lee 2021, Lee 2021). Rozy, notably, stands as a pivotal figure for understanding the interaction between virtual influencers and young audiences, particularly in the context of her significant impact on a distinctive demographic of young individuals, including learners of a foreign language. This unique group is not only cognizant of the issues surrounding virtual beings but also harbors formed opinions on the matter, which is why it is important to incorporate the analysis of user comments on Rozy within classroom discussions, particularly in English classrooms. The critical focus of these learning environments should be to foster student motivation, encouraging them to engage with contemporary subjects that resonate with them. These studies pinpoint the limitation in her appeal, grounded in the phenomenon of the uncanny valley, which explores the complex relationship between a robot's human resemblance and the emotional responses elicited (Cho and Han 2022, Conti, Gathani and Tricomi 2022, Drenten and Brooks 2020, Rodrigo-Martin, Rodrigo-Martin and Munoz-Sastre 2021). As discourse on virtual influencers matures, it is set to offer a richer understanding of the societal repercussions and nuanced perspectives associated with these digital entities. Incorporating this evolving dialogue into classroom discussions will be vital, fostering a generation adept at navigating and critiquing the ever-changing digital landscape with depth and discernment.

2.2 Content and Language Integrated Learning (CLIL)

CLIL is an educational approach that emerged in Europe in the 1990s with the purpose of effectively teaching English by integrating both content and language instruction (Coyle, Hood and Marsh 2010). Halliwell (1992) defined CLIL as the integration of language and other subjects to enable practical communication, highlighting the advantage of providing learners with real contexts for language use through cross-disciplinary activities.

According to Coyle et al. (2010), CLIL demands cognitive processes from learners, which they refer to as cognition. This implies that CLIL not only offers learners authentic opportunities for communication but also aids in cognitive development. Evidence from Europe and Asia reveals positive effects on learners' English communication skills and cognitive abilities after the implementation of CLIL programs (Arribas 2016, Lee and Chang 2008). Given these benefits, CLIL can be seen as a necessary instructional method in English education in the country.

Differentiating CLIL from Content-Based Instruction (CBI), Figure 1 illustrates that CBI places a stronger emphasis on language education, while Immersion predominantly focuses on content. In contrast, CLIL integrates both content and language instruction without favoring one over the other.

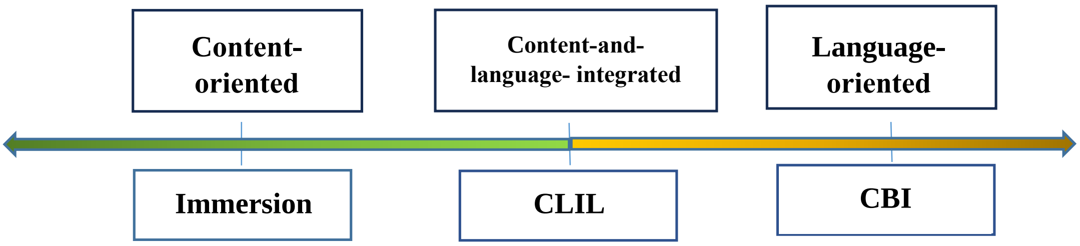


Figure 1. The Comparison of Immersion, CLIL and CBI (Adapted from Kim 2021, p.21)

Coyle et al. (2010) similarly note that Immersion from Canada, CBI from America, and CLIL from Europe are dual-focused pedagogical approaches originating from various regions. However, all these approaches converge in integrating content and language, with CLIL showing distinct characteristics.

The Immersion program teaches all subjects in a second language without explicitly teaching that language, while CBI concentrates on second language learning to bridge subjects and the target language for students with limited English proficiency. CLIL, conversely, prioritizes concurrent content and language instruction without implicit preference for either aspect (Coyle 2007). This distinction is illustrated in Figure 1. Although CLIL and CBI both integrate content and language learning objectives, CLIL is better suited for English as a Foreign Language (EFL) settings and benefits students in terms of native and foreign language as well as content learning. Furthermore, translanguaging strategies are applicable in the CLIL context, using the learners' mother tongue as a tool to learn academic English (Tsou 2018).

In CLIL-based education, it might be proposed to incorporate edtech strategies, specifically by utilizing a word list developed through Python libraries and NLP. This tailored list could potentially familiarize learners with relevant vocabulary. By integrating this list into edtech platforms, educators might have an avenue to personalize quizzes and introduce interactive tools such as flashcards and vocabulary games. This approach could offer a richer learning experience, possibly aiding in assessments, fostering collaborative discussions, and ensuring students engage with crucial vocabulary, pointing towards a data-driven approach in language instruction.

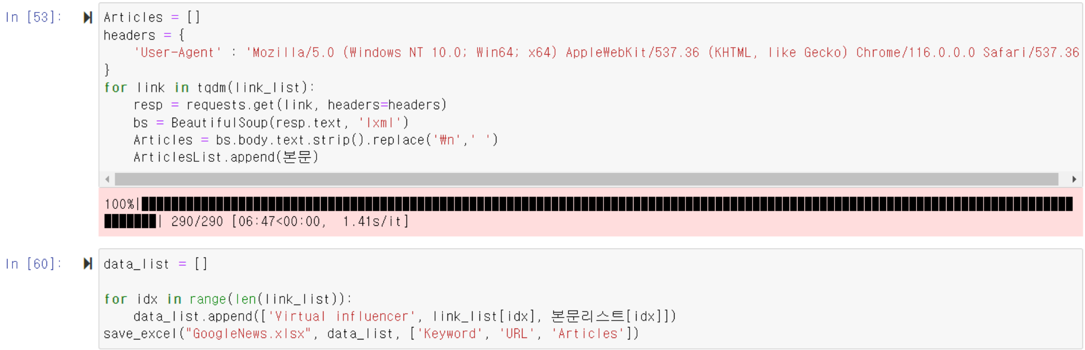
3. Methods

3.1 Data Collection

This study employed two different data collection techniques. The first data collection of comments was called over the span of a year, beginning on July 1, 2021, when Rozy emerged as a virtual influencer. The primary objective of the whole data collection process was to help foster a dynamic environment for English-learning students to critically engage with and discuss the overall landscape of virtual influencers in their own language, which was to ignite interest in English-learning students through Korean comments concerning virtual influencers. The researchers conducted a thorough data collection process, analyzing user comments to spur engagement and reflection on virtual humans' topics. A substantial dataset was gathered from various sources, including 2,375 comments on 151 Naver articles, 1,999 comments on 50 YouTube clips, and 14,253 comments on 287 Instagram feeds using the keyword "gasang ingan Rozy," translating to "the virtual influencer, Rozy." To facilitate this, the NetMiner software and the SNS Collector were utilized for social network and semantic analysis of large datasets, extracting data from platforms like Facebook, Twitter, YouTube, and Instagram. Due to the individual Open API regulations of these platforms, Naver was incompatible with SNS Collector version 2.5.2 (2021b), necessitating Python for data extraction, while YouTube and Instagram data were gathered through the SNS Collector. The research culminated using NetMiner version 4.4.2C for keyword frequency and semantic network analysis, alongside data visualization, effectively encapsulating the research outcomes.

The second data collection method was aimed at developing an authentic English word list to aid students in enhancing their speaking and English composition skills during English classes. To achieve this, the researchers utilized various Python libraries to extract and analyze real-world language usage from platforms like Quora, Google News, and YouTube. This approach was intended to furnish students with a resource that could help them comprehend and adapt to the nuanced English usage in different contexts. Using Python libraries such as Beautiful Soup and NLTK, a considerable volume of text data was extracted from these platforms, forming the basis of the authentic word list. This list would not only facilitate a more realistic approach to English learning but also assist students in distinguishing the varied nuances in language across different mediums, consequently helping to improve their English communication skills. In this way, for the second data collection process, researchers utilized technologies to create a resource for genuine language use, encouraging a more profound understanding of English language intricacies in real-world settings.

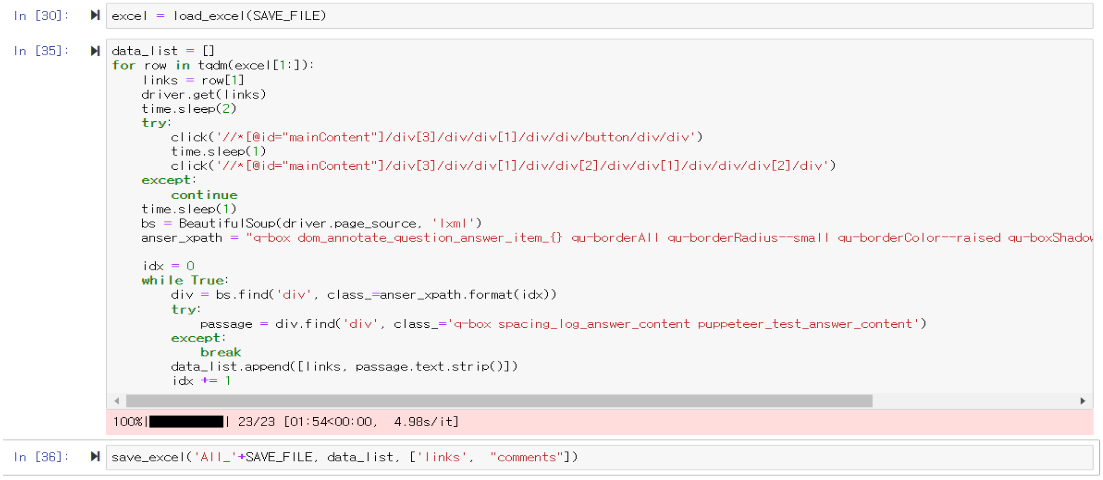
As shown in Figure 2, in order to collect data from Google News articles regarding 'Virtual Influencers,' the Python scripts utilize libraries such as Selenium WebDriver, openpyxl, and others can be imported to facilitate browser automation and data handling. Following this, two functions named get_driver and save_excel are defined, with the former initializing a web driver instance with specific browser settings, and the latter for the process of saving data into an Excel file. Next, the script opens up the web driver and directs it to Google News' search page, where it conducts a search query for "Virtual Influencer." Consequently, a loop commences, iterating 30 times to go through 30 pages of search results. Within each iteration, the script leverages BeautifulSoup to parse the page source and accumulate article URLs into a list named link_list. Additionally, the script navigates to subsequent search result pages by simulating a click on the "Next" button. A 2-second delay is integrated into each loop cycle to avert server overloads by moderating the request frequency. Ultimately, the script gathers a collection of URLs, prepared to be saved for further analysis.



```
In [53]:  Articles = []
          headers = {
              'User-Agent' : 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/116.0.0.0 Safari/537.36
          }
          for link in tqdm(link_list):
              resp = requests.get(link, headers=headers)
              bs = BeautifulSoup(resp.text, 'lxml')
              Articles = bs.body.text.strip().replace('₩n',' ')
              ArticlesList.append(본문)

          100%|██████████████████████████████████████████████████████████████████████
          ███████| 290/290 [06:47<00:00,  1.41s/it]

In [60]:  data_list = []

          for idx in range(len(link_list)):
              data_list.append(['Virtual influencer', link_list[idx], 본문리스트[idx]])
          save_excel("GoogleNews.xlsx", data_list, ['Keyword', 'URL', 'Articles'])
```

Figure 2. Data from Google News Being Collected

As shown in Figure 3 below, to collect each main page and comments of the topic ‘Virtual influencers,’ on Quora, the Python script initiates by importing necessary libraries including Selenium for web automation, BeautifulSoup for web scraping, openpyxl for Excel operations, and tqdm for displaying progress bars. It then defines a set of functions to facilitate browser automation and Excel operations: get_driver to initialize a Selenium WebDriver instance, send_key to send text to webpage elements, click to simulate click actions, get_text to retrieve text from webpage elements, save_excel to save data to an Excel file, and load_excel to retrieve data from an Excel file. Subsequently, the script initializes a WebDriver to navigate to a Quora page concerning virtual influencers, setting up necessary variables for data extraction. A loop then triggers to load a sufficient number of articles by simulating page down key presses and periodically checking the number of loaded articles until it meets a specified minimum. Following this, another loop iterates over each article on the loaded page, attempting to expand them for more detail. After that, the script extracts initial data such as article titles, links, and contents, storing them in a list which is then saved to an Excel file using save_excel. Then, the script reloads this data from the Excel file and iterates over each link to extract further details including comments, which is subsequently saved to a new Excel file, thereby compiling a rich dataset ready for analysis. The entire operation exemplifies a powerful approach to web data extraction leveraging Python script for the data collection.



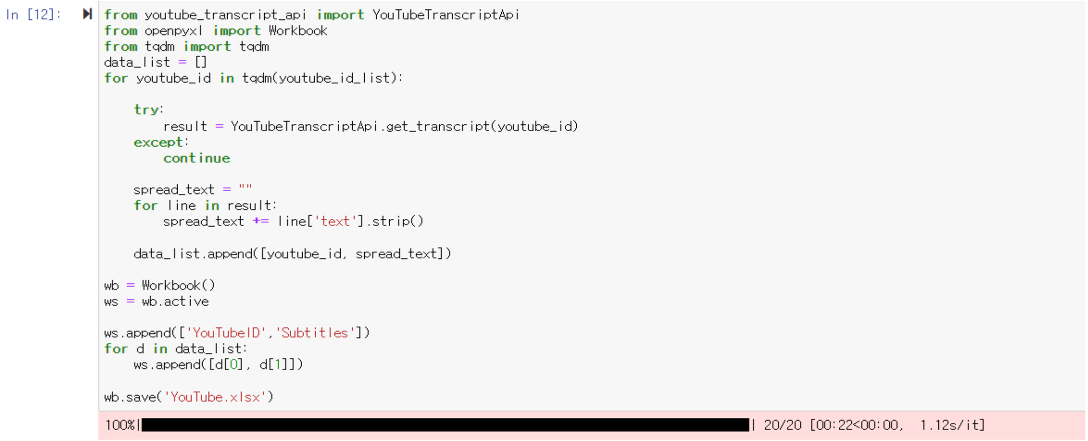
```
In [30]:   excel = load_excel(SAVE_FILE)

In [35]:   data_list = []
           for row in tqdm(excel[1:]):
               links = row[1]
               driver.get(links)
               time.sleep(2)
               try:
                   click('//*[@id="mainContent"]/div[3]/div/div[1]/div/div/button/div/div')
                   time.sleep(1)
                   click('//*[@id="mainContent"]/div[3]/div/div[1]/div/div[2]/div/div[1]/div/div/div[2]/div')
               except:
                   continue
               time.sleep(1)
               bs = BeautifulSoup(driver.page_source, 'lxml')
               anser_xpath = "q-box dom_annotate_question_answer_item_{} qu-borderAll qu-borderRadius--small qu-borderColor--raised qu-boxShadow

               idx = 0
               while True:
                   div = bs.find('div', class_=anser_xpath.format(idx))
                   try:
                       passage = div.find('div', class_='q-box spacing_log_answer_content puppeteer_test_answer_content')
                   except:
                       break
                   data_list.append([links, passage.text.strip()])
                   idx += 1

           100%|██████████| 23/23 [01:54<00:00,  4.98s/it]

In [36]:   save_excel('All_'+SAVE_FILE, data_list, ['links',  "comments"])
```

Figure 3. Data from Quora Being Collected

For YouTube scripts, as shown in Figure 3, the software program initiates by installing the necessary Python packages, YouTube-transcript-API and YouTube-search-python, to facilitate YouTube data retrieval. It then imports required modules and classes from these packages along with openpyxl for data storage and tqdm for displaying a progress bar. Subsequently, it initializes a YouTube search for videos related to the query ‘AI influencer,’ limiting the results to 100 videos, and extracts the IDs of these videos, storing them in a list. Utilizing a loop, it attempts to fetch the transcripts of each video using their respective IDs, and if successful, aggregates the transcript lines into a single string, which is then paired with the video ID in a data list. The collected data is then stored in an Excel file, with columns denoting the ‘YouTubeID’ and ‘Subtitles.’ Separately, the script also fetches the transcript of a specific video identified by the ID “KsukWAE5C_k,” aggregates the transcript lines, and stores this data in a different Excel file, with columns labeled in Korean, translating to ‘YouTubeID’ and ‘Subtitles.’ The script serves as an automated tool for extracting and storing YouTube video transcripts, aiding in

© 2023 KASELL All rights reserved

the large-scale analysis of video content on the platform. The Python scripts[1] can be found and used for research and teaching purposes.



```
In [12]:  from youtube_transcript_api import YouTubeTranscriptApi
          from openpyxl import Workbook
          from tqdm import tqdm
          data_list = []
          for youtube_id in tqdm(youtube_id_list):

              try:
                  result = YouTubeTranscriptApi.get_transcript(youtube_id)
              except:
                  continue

              spread_text = ""
              for line in result:
                  spread_text += line['text'].strip()

              data_list.append([youtube_id, spread_text])

          wb = Workbook()
          ws = wb.active

          ws.append(['YouTubeID','Subtitles'])
          for d in data_list:
              ws.append([d[0], d[1]])

          wb.save('YouTube.xlsx')

          100%|████████████████████████████████████████████| 20/20 [00:22<00:00,  1.12s/it]
```

Figure 4. Data from YouTube Being Collected

3.2 Data Preprocessing Procedure

Upon data collection, a data preprocessing procedure was executed in order to increase the accuracy of data analysis by deleting uninterpretable words and collocating meaningful word sequences. To be more specific, 1) the parts of speech of the keywords were set to extract only “nouns” and “adjectives” from the selected texts, 2) compound words such as “artificial intelligence” were set to treat as a single word, 3) words like “Artificial Intelligence,” “artificial intelligence,” and “AI” were set to be dealt with as a single word, and 4) special characters, symbols, and numbers were excluded. Then, the data was analyzed through NetMiner 4.4.2C.

For the data collected from Quora, Google News, and YouTube had also been preprocessed by using Python scripts. Several libraries such as Selenium, Tqdm, and Openpyxl were initially imported to facilitate web scraping and data storage processes. The Python script initiates by importing the necessary modules to facilitate the tallying of word occurrences, and pandas to manage data effectively. Following this, a function named make_word_frequency is defined to calculate the frequency of words within a given text; it processes the text to remove punctuation and numeric values, normalizes it to lowercase, and utilizes Counter to enumerate the frequency of each word, storing this data in a pandas DataFrame. The script then processes two separate text files in two blocks of similar structure: initially, it reads data from a text file and undergoes a series of preprocessing steps including the removal of certain characters, standardization of quotation marks, and segmentation of text into lines based on specific delimiters; subsequently, it filters and cleans individual lines based on a set of criteria, aggregates them into a singular text string, and writes this preprocessed text into a new file. Parallelly, the make_word_frequency function is invoked to compute the frequency of words in the preprocessed text, the results of which are saved into a CSV file. This procedure is mirrored for the second text file with slight variations to adapt to its unique structure, resulting in another set of preprocessed text and word frequency CSV files. The script exemplifies a structured approach to textual data analysis, showcasing the utility of Python in data science and text analytics tasks.

[1] The following is the link for the Python scripts that anyone interested can refer to: https://github.com/MARTINCAN94/KASELL.

© 2023 KASELL All rights reserved

3.3 Data Analysis

For data analysis, keyword network analysis was applied to elucidate how certain keywords frequently appear together, a critical aspect in comprehending the subject matter at hand. This analytical approach facilitated a deeper insight into both individual node attributes and the holistic properties of the network, using metrics such as “degree,” representing the word's localized connections, and “density” and “centrality,” which delineate broader network characteristics. Furthermore, the centrality analysis, encapsulating ‘degree centrality,’ ‘betweenness centrality,’ and ‘closeness centrality,’ was vital in illustrating the central positioning of nodes within the network structure. Specifically, ‘degree centrality’ was leveraged, calculating centrality based on the total number of edges linked to a single node, to gauge the core influence points within the network. Consistent with methodologies adopted in past research (Kwon 2019, Kwon 2022), the findings of the data analysis were conveyed visually, utilizing word clouds and 2-D spring maps grounded on degree centrality in the word network. This visualization technique allowed for an accessible understanding of the prevailing trends and focal discussion points regarding the phenomenon of virtual influencers, fostering a comprehensive grasp of the topic among students.

4. Results

The study attempted to analyze the Korean social media users’ perceptions of “virtual influencer Rozy” based on the comments on Naver news articles, YouTube video clips, and Instagram feeds on the virtual figure to stimulus. To this end, the researchers conducted keyword frequency analysis followed by network analysis and co-occurrence keywords analysis.

Table 1. Frequency of Top 10 Keywords (by Platform)

Platform	Naver		YouTube		Instagram	
Rank	Keyword	Frequency	Keyword	Frequency	Keyword	Frequency
1	virtual human (*gasang-ingan*)	308	**person (*saram*)**	391	Rozy	642
2	**person (*saram*)**	279	virtual *(gasang)*	263	influencer	535
3	none (*up-da*)	268	**same (*gat-da*)**	224	gram (*geuraem*)	482
4	**same (*gat-da*)**	242	**good (*joh-tah*)**	192	virtual *(gasang)*	317
5	**human (ingan)**	185	none (*up-da*)	183	virtual	270
6	**advertisement (*gwang-go*)**	185	face (*eolgul*)	167	star	252
7	virtual *(gasang)*	184	Rozy	155	gratitude (*gamsa*)	248
8	**like (*joh-ta*)**	177	**advertisement (*gwang-go*)**	152	Oh Rozy (*oh-roji*)	177
9	celebrity (*yeonyein*)	175	virtual human (*gasang-ingan*)	150	**daily life (*il-sang*)**	157
10	article (*gisa*)	165	**human (*ingan*)**	124	**daily(*deilli*)**	143

Table 1 demonstrates ten most frequent words that appear in user comments on the three platforms: Naver, YouTube, and Instagram. The term “person (*saram*)” is ranked at the top on both Naver and YouTube but at the 20th on Instagram. The same can be said for the term “human (*ingan*)” which is synonymous with “person.” While “human” is the 5th and 10th most frequently used words in Naver and YouTube comments, it is non-existent in the

© 2023 KASELL All rights reserved

top 100th on Instagram. Another noteworthy aspect is that verbs such as “like (*joh-ta*)” and “is similar to (*gat-da*)” appear within the top 10th most frequent words in both Naver and YouTube comments, they are far less frequently used in Instagram comments. A more in-depth reading of these differences will be provided later in Table 5, which demonstrates co-occurrence network analysis. Furthermore, both Naver and YouTube comments list “advertisement (*gwang-go*)” respectively at the 6th and 8th, and money-related words like “money (*don*)”—24th on YouTube—and “advertisement cost (*gwang-go-bi*)”—50th on Naver—make frequent appearance, Instagram comments frequently contain fashion-related terms that do not appear in the other two platforms. Those words include “daily life (*il-sang*),” “daily (*dae-il-li*),” “fashion,” “selfie,” “model,” and “outfit (*codi*).” The words are reflections of Rozy’s Instagram captions, a brief description that appears underneath a photo on Instagram, which means that they are repetitions of and replies to Rozy’s words rather than the users’ own thoughts. Finally, combined word frequency of Naver, YouTube, and Instagram comments is visualized via word cloud as in Figure 5.

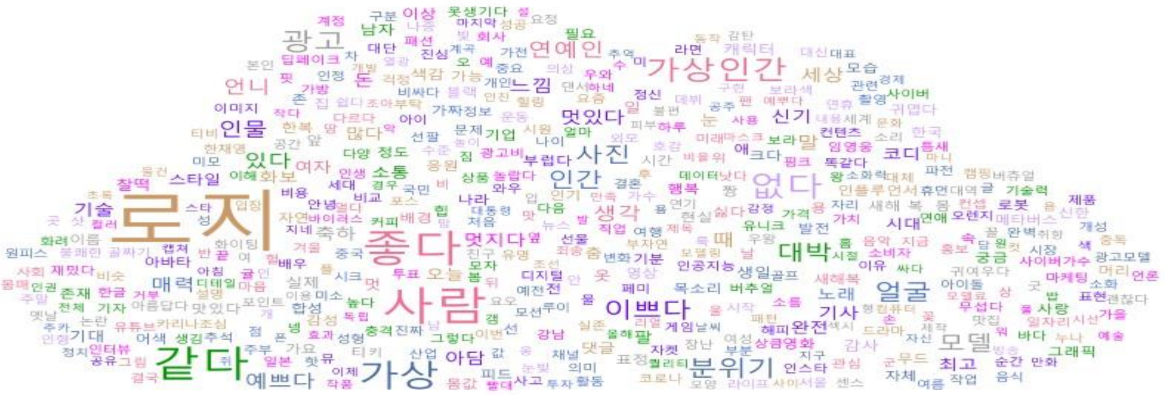


Figure 5. Word Cloud of Keywords (Combined Data from Naver, YouTube, and Instagram)

There are three prominent themes in Figure 5: technology, Rozy’s appearance, and advertisements. The technology-related words from the table are “fascinating (*sin-gi*/29th),” “technology (*gi-sool*/33rd),” “weird (*ih-sang*/55th),” “video (*young-sang*/85th),” and “fake (*ga-jja*/93rd).” They are a mixture of words with both positive and negative connotations and a further analysis of those words will be provided in Table 3, the co-occurrence analysis. The second theme is Rozy’s appearance. The related words are “pretty (*ih-ppeu-da*/10th,, *ye-ppeu-da*/21st),” “face (*ul-gul*/11th),” “celebrity (*yeon-ye-in*/16th),” “attractiveness (*mae-ryeok*/23rd),” and “cool (*meot-ji-da*/28th).” The words associated with Rozy’s looks tend to be positive. Furthermore, words like “eyes (*noon*/40th)” and “arms (*pal*/41th)” also signal high interest in Rozy’s physical appearance. The third theme is related to advertisements. As examined in Table 1, words like “outfit (*codi*/44th)” and “clothes (*ott*/47th),” along with words like “envy (*boo-rup-da*/80th),” indicate users’ interest in and admiration of Rozy’s fashion.

The keywords with the highest degree centrality in the three platforms are in the order as shown in the Table 2.

© 2023 KASELL All rights reserved

Degree centrality is "a simple count of the total number of connections linked to a vertex" (Michigan, Shneiderman, and Smith 2010, p. 71). It is a crude measuring of popularity, meaning that it counts how many edges, that is connections, a given word has to other words. Therefore, it is a simple way of assessing a word's prominence in a given text.

Table 2. Degree Centrality of Top 10 Keywords (by Platform)

Platform / Rank	Naver		YouTube		Instagram	
	Keyword	Centrality	Keyword	Centrality	Keyword	Centrality
1	**there is no (*up-da*)**	0.138158	**person (*saram*)**	0.158921	gram (*geuraem*)	0.10084
2	**virtual human (*gasang-ingan*)**	0.131579	**there is no (*up-da*)**	0.110945	**Rozy**	0.079832
3	**person (*saram*)**	0.131579	**virtual (*gasang*)**	0.089955	daily routine (*ilsang*)	0.077731
4	**is similar to (*gat-da*)**	0.110197	**is similar to (*gat-da*)**	0.088456	star (*star*)	0.073529
5	**celebrity (*yeon-ye-in*)**	0.085526	**like (*joh-ta*)**	0.082459	Fashion (*paesyeon*)	0.067227
6	**human (*ingan*)**	0.082237	**virtual human (*gasang-ingan*)**	0.067466	trip (*yeohaeng*)	0.065126
7	**virtual (*gasang*)**	0.077303	**face (*eol-gul*)**	0.064468	photo shoot (*chwaryeong*)	0.054622
8	**advertisement (*gwang-go*)**	0.067434	**model**	0.05997	styling (*kodi*)	0.054622
9	**good (*jota*)**	0.064145	**think (*saeng-gak*)**	0.055472	daily(*deilli*)	0.048319
10	article (*gisa*)	0.060855	**Rozy (*roji*)**	0.049475	gratitude (*gamsa*)	0.044118

Table 2 shows the words' degree centrality from Naver, YouTube, and Instagram, and it repeats some of the observations from Table 1. Comments from Naver and YouTube share fifteen out of top twenty words with the highest degree centrality: "there is no (*up-da*)," "virtual human (*gasang-ingan*)," "person (*saram*)," "is similar to (*gat-da*)," "celebrity (*yeon-ye-in*)," "human (*ingan*)," "virtual (*gasang*)," "advertisement (*gwang-go*)," "like (*joh-ta*)," "Rozy," "face (*eol-gul*)," "model," "there exist (*it-da*)," "think (*saeng-gak*)," and "Adam." Instagram comments, on the other hand, only share four keywords with the other two platform comments: "Rozy," "model," "like (*joh-ta*)," and "person (*saram*)." As with the Table 1's analysis, Instagram comments are replies to Rozy's original captions and repeats many of Rozy's own words. This reaffirms an earlier observation that Instagram user comments differ from the other two platform users in that they largely approve and reflect what Rozy says in her captions.

© 2023 KASELL All rights reserved

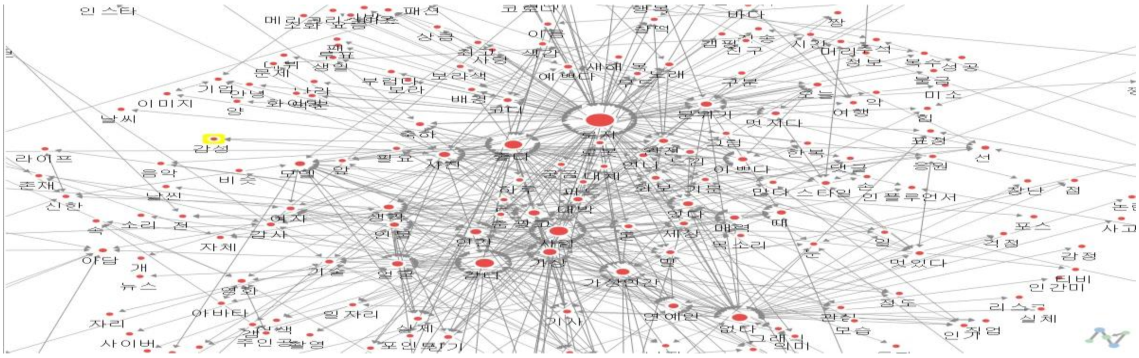


Figure 6. Degree Centrality of Keywords
(Combined Data from Naver, YouTube, and Instagram)

Figure 6 shows combined degree centrality of the three platforms' comments. The list repeats most of the words listed in Table 2. The words in the top twenty, such as "good (*joh-ta*)," "pretty (*i-ppeu-da*)," and "pretty (*ye-ppeu-da*)" indicate positive reception of Rozy. Upon conducting network analysis, the study examined word pairs with strong connection strength based on the frequency of the keywords that occur simultaneously (See Table 3).

Table 3. Top 10 Co-occurrent Keywords (Combined Data from Naver, YouTube, and Instagram)

Platform / Rank	Naver			YouTube			Instagram		
	Source	Target	Weight	Source	Target	Weight	Source	Target	Weight
1	virtual *(gasang)*	human (ingan)	46	virtual (gasang)	character (Inmul)	82	virtual	influencer	486
2	virtual *(gasang)*	character (*Inmul*)	30	person (saram)	face (eolgul)	43	Rozy	Oh Rozy (*oh-roji*)	349
3	**interest *(gwansim)***	**None *(up-da)***	23	**same (gat-da)**	**human (saram)**	36	gram (*geuraem*)	star	307
4	virtual *(gasang)*	model (model)	22	virtual (gasang)	face (eolgul)	33	virtual (gasang)	influencer	290
5	same (gat-da)	person (saram)	21	virtual (gasang)	human (ingan)	30	gram (*geuraem*)	Rozy	288
6	person (saram)	good (jota)	21	virtual (gasang)	model (model)	27	virtual (gasang)	virtual	247
7	**person (saram)**	**none (up-da)**	20	person (saram)	reality (silje)	25	virtual (gasang)	Rozy	241
8	man (nam ja)	woman (yeo ja)	19	technology (gisul)	develop-ment (balgeon)	25	virtual (gasang)	gram (*geuraem*)	205
9	virtual human (gasang ingan)	advertisement (gwang-go)	19	body (mom)	person (saram)	23	gram (*geuraem*)	influencer	205
10	person (saram)	face (eolgul)	17	appliances (gajeon)	house-wife (jubu)	20	gram (*geuraem*)	Oh Rozy (*oh-roji*)	194

Table 3 demonstrates co-occurrence pair of frequently used words in the three platforms' comments. This table makes intelligible what have been discussed so far. Previous discussions showed that verb phrases like "there is no (*up-da*)," "is similar to (*gat-da*)," and "like (*jot-ta*)" locate at the top of both frequency and degree centrality

© 2023 KASELL All rights reserved

measures. Table 3 explains subjects and objects that are related to those verb phrases. Naver comments show the following words' co-occurrence with "there is no (*up-da*)": "interest (*gwansim*)," "person (*saram*)" "work (*il*)," "human (*ingan*)," "necessity (*pil-yo*)," "scandal (*sago*)," and "attractiveness (*mae-ryeok*)." YouTube comments overlap three of these words: "person (*saram*)" "work (*il*)," and "necessity (*pil-yo*)." Instagram comments, on the other hand, share none. This illuminates two things. One is that comments' characteristics differ according to the platforms. While Naver and YouTube comments discussed the phenomenon of virtual influencer in general, Instagram comments largely reflected Rozy's words in her posts and captions. The other is, Naver and YouTube comments are more critical toward Rozy and the phenomenon of virtual influencer, since Naver and YouTube's frequent comments expressed that "they were not interested in" and "do not need" virtual influencers.

Besides, "interest (*gwansim*)" and "there is no (*up-da*)" pair which is 3rd on the Naver list, Naver also lists "look (*ggol*)" and "hate (*sil-ta*)" pair as the 18th highest weighted. In Korean language, "*ggol bogi sil-ta*" is an expression that means one hates the sight of something that they do not wish to look at. This negative emotion reaffirms previous finding that Naver comment users' disapproving opinion toward Rozy as explained with "there is not (*up-da*)" and "interest (*gwansim*)"/"necessity (*pil-yo*)" pairs. YouTube, on the other hand, involves comments that discuss technological development. The "technology (*gi-sul*)" and "development (*bal-jeon*)" pair shows 8th highest weight and related word pairs such as "face (e*ol-gul*)" and "photoshop (*hap-seong*)"; "line (*seon*)" and "dance (*choom*)"; "motion" and "capture"; and "technology (*gi-sul*)" and "like (*joh-ta*)" indicate high interest in technological capability that makes Rozy's facial expression and dance movement appear natural.

Additionally, from this analysis issues regarding public resentment towards celebrity wealth and the "uncanny valley" theory can be discussed in a classroom. There seems to be a noticeable sense of "relative deprivation" among the younger generation in Korea towards celebrity wealth, as indicated by various studies and user comments on the Naver platform. They perceive the wealth and lifestyle of celebrities as undeserved, fostering resentment and a desire for AI influencers like Rozy who might reduce product advertising costs. This sentiment, rooted in existing economic insecurities among younger generations, could potentially be utilized as a discussion point in CLIL (Content and Language Integrated Learning) classes to explore societal trends, economic disparities, and the evolving nature of celebrity culture and advertisement strategies in Korea. Another point of discussion can be the concept of the "uncanny valley," initially proposed by Mori (2012), in relation to virtual influencers like Rozy. While there are mixed responses to Rozy's human resemblance, the discussion tends to lean more towards concerns about the virtual influencer phenomenon rather than Rozy's uncanny appearance per se. It was noted that negative responses are relatively minimal and that many users show affection towards Rozy, indicating a potential shift in perception of virtual influencers. This topic can be introduced in CLIL classes to foster discussions about the advancement of technology, the changing boundaries between humans and AI, and its impact on society and media consumption.

In order to generate a vocabulary list, authentic data were gathered from platforms such as Google News, YouTube, and Quora, following the procedures outlined in the methodology segment of this research. The 'AntConcWordProfiler' software tool was utilized to identify the keywords, which were then compiled into the chart illustrated below (Table 4).

© 2023 KASELL All rights reserved

Table 4. Keywords from Google News, YouTube, and Quora

Platform / Rank	Google News		YouTube		Quora	
	Keyword	Frequency	Keyword	Frequency	Keyword	Frequency
1	influencer	4557	AI	218	influencer	154
2	brand	1825	influencer	109	VR	39
3	digital	1019	GPT	26	AI	27
4	AI	889	Swap	25	brand	26
5	tech	486	fake	23	digital	23
6	Miquela	433	ID	22	Imma	17
7	metaverse	433	Miquela	21	platform	17
8	platform	418	Discord	14	authenticity	14
9	engagement	298	personality	13	personality	12
10	retail	278	digital	12	avatar	7

In the Google domain, the spotlight is on the “influencer (4557)” culture, melding with “brand (1825)” narratives and the reality of the “digital (1019)” era. This interactive narrative, encapsulating AI (889) breakthroughs and emerging “tech (486)” trends, is led by figures like “Miquela (433),” who dominate the discussions about the “metaverse (433).” The “platform (418)” observes a dynamic digital “engagement (298),” where individuals immerse in enriching interactions with virtual entities, significantly influencing the “retail (278)” sectors. Within this, names such as “Kyra (240),” “avatar (177),” and concepts like “copyright (129)” create ripples. Here, the role of individuals like “Maya (124)” and “Kami (119)” becomes pivotal, as conversations revolve around the latest “launch (167)” events and “debut (117)” stories. The ecosystem also thrives on discussions concerning “gen (109)” perspectives, “CGI (105)” developments, personalities like “Imma (104),” and insights from various “CEO (104)” figures, all encased within a thriving digital community fostered by “Brud (94).”

Transitioning to YouTube, a platform where “AI (218)” initiates the conversation, followed by a rich exploration of the “influencer (109)” sphere. Topics like “GPT (26)” generate substantial interest, leading to a fascination with “swap (25)” technologies—technologies that open doors to experiences such as face swapping or ID swapping, providing learners an opportunity to grasp new concepts within context. This space fosters an atmosphere of critique and analysis, touching upon the “fake (23)” personas phenomenon and “ID (22)” validation processes. Users actively engage with platforms like “Discord (14),” where discussions surround digital “personality (13)” and “engagement (6)” with digital entities like “Miquela (21)” take center stage. This dynamic space fosters an environment where “experiential (5)” narratives thrive, with an emphasis on user “feedback (5)” and the creation of “fictional (5)” personas, exploring facets of digital existence ranging from the “Adam (4)” concept to “Asia (4)” focused dialogues, evaluating the “flawless (4)” virtual representation and the “drawback (4)” of digital advancements, all the while cherishing the “genuine (4)” interactions in the digital space.

Lastly, the discourses on Quora unravel a multifaceted narrative encompassing “influencer (154)” dialogues and burgeoning “VR (39)” trends. Conversations on “AI (27)” intertwine with “brand (26)” discussions and “digital (23)” advancements. The spotlight also falls on figures like “Imma (17),” scrutinized within the discourse revolving around “platform (17)” credibility and “authenticity (14).” The dialogues here delve deeper, dissecting the “personality (12)” of virtual beings, the evolution of “avatar (7)” dynamics. Discussions here sprawl into concerns of “scam (7)” scenarios and implications surrounding “brand (7)” imaging. Furthermore, topics such as “render (7),” “generator (7),” “unreal (6),” “chatbot (6),” and “icon (6)” not only add depth to the conversation but offer a rich ground for exploration and learning. Thus, the wordlist can be used to stir up learner engagement for learning and using words with authentic materials for speaking and writing activities in a CLIL classroom based on the discussion points derived from the analysis of domestic sentiments.

5. Discussions and Conclusion

In the expansive realm of digital discourse surrounding virtual influencers, collections of data extracted from various platforms reveal profound people's insights. This study unravels the intricate web of narratives and public perceptions stemming from diverse sources-from Korean comments on the Naver platform and various YouTube and Instagram comment sections to YouTube scripts and English news articles on Google News and Quora responses.

This study analyzed user comments in three popular social media platforms in Korea-Naver, YouTube, and Instagram—by using text-mining method. Its aim was to elucidate what opinions and feelings people have toward Rozy specifically and about virtual influencers generally. As a relatively new phenomenon but one with high potential of growth, Rozy or virtual influencer is an important subject of study. This study found stark differences in contents and sentiments in user comments in the three platforms and found the causes in its functions and varying levels of user anonymity. Naver news articles, one with the highest level of user anonymity, attracted the most critical comments about Rozy; YouTube, a video-sharing platform, attracted the most CGI technology-related comments. On the other hand, Instagram, with the lowest level of user anonymity, attracted the most favorable comments. Furthermore, this study also found that users were feeling a sense of "relative deprivation" at celebrities who, in their opinion, were living luxurious lifestyles. Users expressed such resentment by using Rozy as a proxy through which they can make even with celebrities. In other words, celebrities will no longer be able to have it so easy in their lives as virtual influencers will gradually take over their jobs. Last finding was that unlike previous studies about virtual influencers and robots, "uncanny valley" did not have significant impact on public perceptions about Rozy.

Shifting the focus to the English wordlist created from authentic materials derived from YouTube scripts, English news articles on Google News, and responses on Quora, the analysis of the English corpus from Google News underscores a pronounced emphasis on influencer culture, which is closely connected with narratives around brand identity in the digital domain. It is apparent that there is an increasing interest in technological developments, including AI and emerging tech trends, fostering vibrant discussions centered on the concept of the "metaverse." Noteworthy virtual influencers like Miquela are leading the conversations, considerably shaping public perception and engagement across various sectors, including retail.

In contrast, the YouTube platform serves as a fertile ground for discussion where narratives revolving around AI and influencers hold a central position. There is a marked focus of discussion points on technology related topics such as GPT and technology swaps catching significant attention. These encourage users to delve into and critique concepts like fake personas and partake in discussions regarding ID verification processes. These conversations frequently extend to platforms like Discord, providing a lively environment for enhanced digital interactions.

Meanwhile, Quora is experiencing a surge in discussions related to influencers and burgeoning VR trends, intertwining with conversations about brand development and digital advancements. In this space, the discourse deepens, centering on critical analysis encompassing themes of authenticity, personality dynamics, and avatar developments, among others.

Overall, this research underscores the pivotal role platform characteristics play in shaping the discourse, resonating well with Marshall McLuhan's assertion that "the medium is the message." In this scenario, a nuanced approach to data analysis becomes crucial, acknowledging the subtle distinctions between platforms and the content they host to construct a detailed and precise portrait of the swiftly transforming digital narrative landscape.

As research with big data analysis via social media, this study contains some inherent limitations as a natural result of the methodology. First, it requires a cautious interpretation of data analysis since they are based on the indirect information withdrawn from big data, not on the direct information acquired from surveys, interviews, etc. (Ham and Chae, 2012). Second, big data analysis of social network platforms has only recently begun to gain attention in the field of social science, and it has often been limited to exploratory research. Lastly, whether the comments from the selected social media site during a specific period of time accurately represent the public's opinion on the concerned issue is questionable. Additionally, when the amount of data attained from one media platform predominates, its effect on the results of the data analysis may skew the analysis.

Regardless of the above limitations, however, the study is meaningful in that 1) it quantitatively analyzed big data composed of a large amount of non-structured text; 2) it also deals with a timely topic in the current situation while the public interest in virtual influencers is rising in a CLIL classroom; 3) the natural and detailed expressions of the public in the collected comments from each social media platform contains information that is hard to obtain from surveys or other traditional means of data collection; 4) it clearly reveals the varied characteristics of each social media platform users through media-specific comments on the virtual influencer Rozy; and 5) it showcases the use of authentic materials turned into a wordlist to help learners' engagement.

As researchers, it becomes imperative to recognize and appreciate these intricate differences between platforms and the type of content (be it the main body of text, comments, captions, or others) they host. A nuanced approach to data collection thus becomes essential, where not just the platform's unique characteristics, but also the specific type of content within the platform needs to be considered, to sketch a more accurate and comprehensive picture of the evolving digital narrative landscape.

References

Alexa. 2022. *Top sites in South Korea.* Available online at https://www.alexa.com/siteinfo.

Ahn, H. S. and Lee, T. H. Lee. 2021. Virtual influencers kept busy as jobs keep coming in. *Korea Joongang Daily*, August 14, 2021. Available online at https://koreajoongangdaily.joins.com/2021/08/14/business/industry/virtualinfluencer/20210814070114833.html.

Arribas, M. 2016. Analysing a whole CLIL school: Students' attitudes, motivation, and receptive vocabulary outcomes. *Latin American Journal of Content & Language Integrated Learning* 9(2), 267-292.

Baek, J. 2020. 'Hateful comments not allowed'...Naver discloses comment history. *Seoul Economist*, March 18, 2020. Available online at https://www.sedaily.com/NewsView/1Z08VJHHQ3/GD0503.

Black, D. 2008. The virtual ideal: Virtual idols, cute technology and unclean biology. *Continuum* 22(1), 37-50.

Brunjes, K. 2022. Age range by generation. *Beresford Research* (blog). Available online at https://www.beresfordresearch.com/age-range-by-generation/.

Chase, K. and K. Johannsen. 2019. *World English 2 with my world English*. Boston: Heinle ELT Publishing.

Cho, D. and G. Han. 2022. How virtual influencer characteristics affect purchase intention: Focusing on uncanny valley theory. *The Korea Journal* of *Advertising and Public Relations* 24(3), 135–69.

Choi, J. B. and M. Lee. 2017. News content consumption analysis of news consumers in the era of new media. *The Journal of the Korea Contents Association* 17(2), 207-218.

Conti, M., J. Gathani and P. P. Tricomi. 2022. Virtual influencers in online social media. *IEEE Communications Magazine* 60(August), 1-13.

Copeland, C. 2021. Issue brief no. 548. Comparing the financial wellbeing of baby boom, generation X, and

© 2023 KASELL All rights reserved

millennial families. *EBRI.* Available online at https://www.ebri.org/docs/default-source/ebri-issue.

Colye, D. 2007. Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism* 10(5), 543-562.

Coyle, D., P. Hood. and D. Marsh. 2010. *Content and language integrated learning.* New York: Cambridge University Press.

Drenten, J. and G. Brooks. 2020. Celebrity 2.0: Lil Miquela and the rise of a virtual star system. *Feminist Media Studies* 20(8), 1319-1323.

Halliwell, S. 1992. *Teaching English in the primary classroom*. London: Longman.

Ham, Y. and S.-B. Chae. 2012. *Big data changes management.* Seoul: Samsung Global Research.

Hiort, A. 2022. Yes, virtual influencers are taking jobs. *Virtual Humans*, February 15, 2022. Available online at https://www.virtualhumans.org/article/yes-virtual-influencers-are-taking-jobs.

Hong, J. I. 2021. After displaying the profile picture of the author in the ‘comment on Naver, malicious comments decreased by 16%. *Yonhap News*, July 8, 2021. Available online at https://www.yna.co.kr/view/AKR20210708131500017.

Hwang, S. and M. C. Lee. 2021. Analysis of the value change of virtual influencers as seen in the press and social media using text mining. *The Korean Journal of Advertising and Public Relations* 23(4), 265–99.

Jang, Y. 2017. Global economic optimism index. *Yonhap News*, February 6, 2017. Available online at https://www.yna.co.kr/view/GYH20170206001200044.

Joo, Y.-J. 2018. Portal websites struggle in the YouTube era. *The Kyounghyang Shinmun*, March 8, 2018. Available online at https://news.khan.kr/4shR.

Kim, H. 2021. *A Study on the Effect of Character-Centered CLIL on English Learning and Character Development of EFL Children*, Doctoral dissertation, Hannam University, Daejeon, Korea.

Kobayashi, H. and T. Taguchi. 2019. Virtual idol Hatsune Miku: Case study of new production/consumption phenomena generated by network effects in Japan’s online environment. *Markets, Globalization & Development Review* 3(4).

Kwon, E.-Y. 2019. Public perception of ‘early English education’ through an analysis of social media big data: Focusing on YouTube. *Korean Journal of English Language and Linguistics* 19(4), 858-879.

Kwon, E.-Y. 2022. Research trends of domestic studies on teaching tools in English language teaching. *Korean Journal of English Language and Linguistics* 22, 637-660.

Lee, B. and K. Chang. 2008. An overview of content language integrated learning in Asian countries. *Studies in English Education* 13(2), 166-184.

Lee, H. 2021. The effects of virtual influencer marketing on consumers' brand attitude and purchase intention: Based on the unpleasant valley effect and the recognized characteristics of virtual influencer 'Rozy.' *2021 Korea Speech, Media & Communication Association Conference*, 134-47.

Lee, H. and J. Ma. 2022. Virtual influencer ‘ROZY’: Is the attitude towards ‘her’ same for everybody? Segmentation approach based on virtual model attitudes. *Korea Logistics Review* 32(3), 43-55.

Lee, J. and C. Lee. 2022. Formative characteristics of virtual influencer Imma’s fashion. *Illustration Forum* 23(72), 85-96.

Lee, J. K. and Y. Choi. 2018. The effect of SNS type on brand loyalty and attitude toward AD: Focusing on the mediating effects of brand personality. *Journal of Business Research* 33(2), 75-96.

Lee, S. 2021. *The Effect of Virtual Fashion Influencer's Presence on Evaluation Attributes and User Responses*. Master’s thesis, Seoul National University, Seoul, Korea.

McLuhan, M. and L. H. Lapham. 1994. *Understanding media: The extensions of man. (Reprint ed.).* Cambridge: The MIT Press.

© 2023 KASELL All rights reserved

McSpadden, K. 2015. You now have a shorter attention span than a goldfish. *Time,* May 14, 2015. Available online at https://time.com/3858309/attention-spans-goldfish/.

Michigan, D. H., B. Shneiderman and M. Smith. 2010. *Analyzing Social Media Networks with NodeXL: Insights from a Connected World (1st ed.).* Burlington, MA: Morgan Kaufmann.

Mori, M. 2012. The uncanny valley [from the field]. Translated by K. F. MacDorman and N. Kageki. *IEEE Robotics & Automation Magazine* 19(2), 98-100.

Park, J. 2022. *Analysis of Virtual Influencers' Attractive Physical Appearance on Social Media.* Master's thesis, Seoul National University, Seoul, Korea.

Park, J. and Y. Koo. 2022. Analysis of SNS content usage types and consumer reactions of virtual influencers - Focusing on the case of virtual influencer 'Rozy.' *Journal of Communication Design* 81, 593-609.

Park, J. 2022. A study on the relationship between virtual influencer attributes, imitation intention, and usage intention. *The Journal of the Convergence on Culture Technology* 8(3), 245-51.

Park, Y., D. Shin, J. Kwon, J. Park, Y. Guo and J. Yoon. 2022. A study on user preference based on the characteristics of virtual influencers. *Design convergence study* 21(2), 1-16.

Rahmi, M. S., N. Rahmat and A. Saleha. 2018. Posthuman in Japanese popular culture: Virtual idol Hatsune Miku. *AICLL: Annual International Conference on Language and Literature* 1(1). 81-86.

Rasmussen, M. 2022. The top 10 virtual influencers in Korea. *Virtual Humans.* Available online at https://www.virtualhumans.org/article/the-top-10-virtual-influencers-in-korea.

Robinson, B. 2020. Towards an ontology and ethics of virtual influencers. *Australasian Journal of Information Systems* 24(June), 1-8.

Rodrigo-Martin, L., I. Rodrigo-Martin and D. Munoz-Sastre. 2021. Virtual influencers as an advertising tool in the promotion of brands and products: Study of the commercial activity of Lil Miquela. *Revista Latina de Comunicación Social* 79, 70–91.

Smith, H. J. and T. F. Pettigrew. 2015. Advances in relative deprivation theory and research. *Social Justice Research* 28(1), 1–6.

The Deloitte Global. 2022. Gen Z and millennial survey. Available online at https://www.deloitte.com/global/en/about/people/social-responsibility/genzmillennialsurvey.html.

Tsou, W. 2018. Implementing content language integrated learning (CLIL) in Taiwan: A review study. *Proceedings of the PIM 8th National and 1st International Conference*, 1-10.

Twenge, J. M., G. N. Martin and B. H. Spitzberg. 2019. Trends in U.S. adolescents' media use, 1976–2016: The rise of digital media, the decline of TV, and the (near) demise of print. *Psychology of Popular Media Culture* 8(4), 329–45.

Walch, K. 2022. The creepy problem killing AI projects. *Forbes,* September 17, 2022. Available online at https://www.forbes.com/sites/cognitiveworld/2022/09/17/the-creepy-problem-killing-ai-projects/.

Yeung, J. and G. Bae. 2022. Forever young, beautiful and scandal-free: The rise of south Korea's virtual influencers. *CNN,* August 16, 2022. Available online at https://www.cnn.com/style/article/south-korea-virtual-influencers-beauty-social-media-intl-hnk-dst/index.html.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

© 2023 KASELL All rights reserved