Korean Journal of English Language and Linguistics, **Vol 26, February 2026, pp. 198-214**
DOI: 10.15738/kjell.26..202602.198

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

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

http://journal.kasell.or.kr

Thematic Patterns in Tertiary-Level ELT Research During the COVID-19 Pandemic in Korea: A Text Mining Analysis

Sojeong Kim (Dongyang Mirae University)

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: November 10, 2025
Revised: January 15, 2026
Accepted: January 15, 2026

Kim, Sojeong
Associate Professor
Department of General Education
Dongyang Mirae University
445 Gyoungin-ro, Guro-gu, Seoul, Korea
Tel: +82-2-2610-5198
Email: sunsj@dongyang.ac.kr

ABSTRACT

Kim, Sojeong. 2026 Thematic patterns in tertiary-level ELT research during the COVID-19 pandemic in Korea: A text mining analysis. *Korean Journal of English Language and Linguistics* 26, 198-214.

Generative AI is rapidly being integrated into English education, marking a new paradigm of innovation. Having experienced the pandemic's abrupt shift in learning environments without adequate preparation, the present moment calls for proactive readiness. South Korea, where enthusiasm for English learning and an advanced ICT infrastructure create a unique context, could offer valuable insights. This study examines thematic patterns in Korean tertiary-level English education research during the COVID-19 pandemic. Text mining was applied to 190 Korea Citation Index (KCI)-indexed journal articles using three analytic methods: frequency analysis, semantic network analysis, and Latent Dirichlet Allocation (LDA) topic modeling. The findings indicate that learner-centered themes emerged as major areas of focus. Frequency analysis showed "learner" in 98.4% of documents, "online" in 91.6%, and "learning" in 76.8%. Semantic network analysis identified "learner" as the central node, most strongly linked with "online" (91.1%) and "learning" (75.8%) and highlighted low-proficiency learners, as "beginner-level" and "beginner-proficiency" pairs ranked among the top five associations. LDA topic modeling identified eleven thematic clusters, with learner satisfaction, engagement, and autonomy emerging alongside instructional design issues such as delivery formats, digital tools, and blended learning. Overall, the results suggest that the central pedagogical concern was not technology but the learner. In this new paradigm, as generative AI transforms educational practice, learners may become active participants who design, regulate, and critically engage with their own learning. Future research should explore how these evolving roles can be effectively supported through pedagogical design.

KEYWORDS

ELT, COVID-19, thematic analysis, text mining, topic modeling, learner-centered instruction

© 2026 KASELL All rights reserved

1. Introduction

We are now entering an era where generative AI is rapidly reshaping not only daily life but also the fundamental structures of education. This shift is prompting a new phase of experimentation and innovation in English Language Teaching (ELT), particularly through the emergence of tools such as ChatGPT (Law 2024, Lo et al. 2024, Rifqiyah et al. 2025). In such times of accelerated innovation, revisiting prior experiences of educational disruption, such as the COVID-19 pandemic, can offer valuable insights for navigating current and future changes.

The COVID-19 pandemic led to an unprecedented global shift from face-to-face to online instruction, referred to as Emergency Remote Teaching (ERT) (Hodges et al. 2020). Unlike pre-planned online learning, ERT required both instructors and learners to adapt rapidly to unfamiliar online learning environments, often without adequate training or infrastructure (Bozkurt and Sharma 2020, Crawford et al. 2020, Mishra et al. 2020). In the context of English language teaching, where communication and interaction are essential, this abrupt transition introduced unique pedagogical and emotional challenges (Moorhouse and Kohnke 2021). Notably, the pandemic gave rise to a unique moment in educational history, when digital instruction was no longer optional but instead became a universal and mandatory mode of teaching and learning. Unlike the pre-pandemic period when digital tools were selectively adopted, the pandemic forced all instructors and students to engage with online platforms, regardless of prior experience. This large-scale, involuntary implementation of educational technologies provides a rare opportunity to examine how pedagogical priorities evolved under crisis conditions.

Recent literature reviews have analyzed this transformation and identified three dominant themes: (1) digital transformation and instructional innovation, (2) learner adaptation and autonomy in online learning, and (3) the adoption of hybrid and blended learning models as future-oriented instructional approaches. Together, these themes highlight how English language teaching research conceptualized educators' and learners' responses to the transition to online instruction.

Although research on the topic has grown, substantial gaps remain, especially regarding underrepresented regions outside the English-speaking or technologically advanced world (Isti'anah and Wijanarka 2023, Mishra et al. 2021). South Korea, in particular, presents a unique context with advanced ICT infrastructure and a strong societal emphasis on English proficiency. In South Korean universities, English-related courses account for approximately 17% of all general education offerings (Korean Council for University Education 2023), underscoring the significant role of English education. In this context, as digital learning environments become increasingly integral to English education, examining South Korea's experience offers valuable insights into broader educational developments.

As technology-enhanced learning once again expands rapidly, this time through generative AI tools, understanding the themes that emerged during the pandemic in a digitally advanced context like Korea may offer foundational insights for future meta-analyses and educational design frameworks. Against this backdrop, the present study investigates how tertiary-level ELT research in South Korea responded to the challenges of the COVID-19 pandemic. By analyzing KCI-indexed journal articles through text mining techniques, the study aims to identify key themes and draw implications for ELT in evolving digital learning environments integrated with generative AI.

© 2026 KASELL All rights reserved

2. Literature Review

2.1 Global Trends in Tertiary ELT during the Pandemic

A range of systematic reviews and bibliometric studies on ELT during the pandemic (e.g., Erarslan 2021, Isti'anah and Wijanarka 2023, Martin et al. 2021, Moorhouse and Kohnke 2021, Soon and Aziz 2022) examined how ELT adapted to this global crisis, shedding light on both the immediate pedagogical responses and shifts in teaching and learning practices. While these studies encompassed different levels of education, their findings provide important insights into tertiary-level ELT, particularly in terms of (1) teachers' instructional innovation (Moorhouse and Kohnke 2021), (2) learner adaptation and autonomy in online learning (Mahyoob 2020), and (3) the adoption of hybrid and blended learning models as future-oriented instructional approaches (Ramalingam et al. 2022).

Recent review studies on higher education during the COVID-19 pandemic have collectively focused on restructuring the pedagogical framework of ELT, ranging from instructor adaptation to specific learning strategies. Martin et al. (2021) examined the challenges faced by ESL instructors during the sudden transition of curricula, identifying critical technical, pedagogical, and psychological difficulties. Their findings underscored the urgent need for strengthening digital teaching competencies and developing standardized online instructional designs to ensure the stability of higher education curricula.

Alongside these instructional considerations, the significance of technological frameworks was further highlighted by Ramalingam et al. (2022). Recognizing that the pandemic environment had fundamentally intensified the importance of technology integration in ESL pedagogy, they conducted a systematic review of blended learning strategies in articles published between 2017 and 2022. Their analysis identified four key themes, with technology-based instruction as the most prominent trend. Notably, they argued that technology-driven approaches are pivotal in addressing issues of language proficiency, motivation, and learner engagement in tertiary-level ESL education.

Specific pedagogical strategies were also validated for their efficacy in enhancing the learning experience. Khreisat (2022) investigated diverse learning strategies through a review of pandemic-era studies, demonstrating that cooperative learning, flipped classrooms, and scaffolding were effective in improving instructional efficacy within EFL environments.

However, several reviews have noted regional and thematic imbalances. Mishra et al. (2021) reported that approximately 80% of research on online distance learning during the pandemic was conducted in just 20% of countries. Similarly, Isti'anah and Wijanarka (2023) identified a lack of cross-national collaboration in ELT research during the pandemic, highlighting the need for increased international joint research.

2.2 ELT Research in Korean University Context: Pre- and Post-Pandemic

In light of the regional imbalances in global ELT research, South Korea could offer valuable perspectives to the international ELT field. Its robust ICT infrastructure and strong societal emphasis on English proficiency enabled active research and experimentation with digital pedagogies during the pandemic. In addition, English-related courses occupy a considerable share of the general education curricula within Korean higher education. Since the late 1990s, various efforts have been made to implement new general English programs, including the expansion of English-medium instruction (EMI), the development of level-based curricula, and the integration of English for Specific Purposes (ESP) (Kim and Lim 2013). These educational innovations, combined with South Korea's robust

© 2026 KASELL All rights reserved

ICT infrastructure, have supported ongoing digitalization and pedagogical innovation in English instruction. Alongside these developments, research trends in English education have also undergone significant changes.

Prior to the pandemic, research in tertiary-level ELT revealed a long-term shift. Ha (2017) identified that, before the 2000s, research largely focused on curriculum development, teaching methods, and assessment, whereas studies after 2000 shifted toward learner-centered approaches and the integration of digital tools. Park (2021a) identified ten major themes in college English education research from 2001 to 2020 using topic modeling. Among these, curriculum and level-differentiated programs (18.67%) ranked first, followed by a relatively even distribution of other themes: learner-centered education (11.67%), teachers' factors (10.67%), textbook analysis (9.67%), learners' affective factors (9.33%), assessment and learning strategies (9.00%), and English vocabulary, grammar, and writing (9.00%). Notably, the study identified learners' affective factors, English for Specific Purposes (ESP) programs, and learner-centered education as rising themes after 2011. Similarly, Park (2021b) found that while the most prominent keywords between 2001 and 2020 were questionnaire, course, program, level-differentiated, and test, there was a discernible upward trend in keywords such as motivation, attitude, perception, and learners' satisfaction from 2010 to 2020.

In post-pandemic research in tertiary-level ELT, Ha (2022) conducted a qualitative exploration of flipped-learning-based college English education, highlighting the practical challenges and pedagogical possibilities during the digital transition. Ha (2022) emphasized the need for comprehensive follow-up research on diverse teaching models and methods for flipped-learning-based college English education in the post-pandemic era, while also highlighting the importance of institutional policy support for technology use and specialized support models for lower-level learners. Lee et al. (2023) observed a thematic shift toward online teaching methods and self-directed learning, moving away from the earlier emphasis on traditional instructional techniques and assessment practices. In this context, there remains a critical need for an in-depth investigation into the thematic patterns that emerged within South Korea's tertiary education throughout the pandemic. The present study aims to fill this gap by identifying the unique research trajectories and pedagogical patterns in university-level English education under pandemic constraints, utilizing text-mining techniques.

2.3 Text Mining Approaches in ELT Research

Text mining is defined as the process of extracting useful information from unstructured textual data through the identification and exploration of patterns (Feldman and Sanger 2007), by applying methods such as text classification and clustering, natural language processing, information retrieval, and text summarization (Ferreira-Mello et al. 2019). Various text mining techniques, including frequency analysis, co-occurrence analysis, bibliometric analysis, centrality analysis, topic modeling, and clustering, have been employed to investigate research trends and thematic structures (see Table 1).

As shown in Table 1, co-occurrence analysis and topic modeling appear to be frequently used techniques. Co-occurrence patterns serve as indicators of conceptual relationships by capturing how terms appear together across documents. For example, Fujii et al. (2022) employed co-occurrence network analysis to examine learner responses, finding that activities such as asking questions, collaborating with peers, and receiving feedback from teachers were more closely associated with higher satisfaction in synchronous learning environments. Topic modeling, another widely adopted technique, allows researchers to uncover latent thematic structures in large text datasets (Alghamdi and Alfalqi 2015, Blei et al. 2003). It is frequently used to map research trends, explore instructional strategies, and categorize educational materials. Chen et al. (2021) identified 15 major research topics and revealed how these topics evolved over time, using structural topic modeling on 1,295 CALL-related articles

© 2026 KASELL All rights reserved

published from 1995 to 2019. Their analysis demonstrated a shift in research interest from traditional topics to more diverse tools and applications. This finding highlights the value of topic modeling for mapping long-term thematic trends in language education research.

Given the unique strengths of each technique, researchers often combine multiple text mining techniques rather than relying on a single technique. Building on this approach, the present study employs frequency analysis, semantic network analysis, and LDA topic modeling to capture thematic dynamics in Korean ELT research during the COVID-19 pandemic.

3. Methodology

3.1 Data Collection and Selection Criteria

This study collected research articles published in the KCI that addressed English teaching and learning at the tertiary level during the COVID-19 pandemic (March 2020–June 2024). Searches were conducted in both Korean and English using combinations of "English teaching" or "English learning" with online-related terms (e.g., "Emergency Remote Teaching (ERT)," "COVID-19," "pandemic," "online," "Zoom," "synchronous," "asynchronous," and "non-face-to-face"). After screening 2,289 retrieved papers, a total of 190 articles were selected for the final dataset. The yearly breakdown is as follows: 21, 75, 51, and 43 articles in 2020, 2021, 2022, and 2023, respectively. Although the search was extended to June 2024, no articles met the inclusion criteria during that period. All selected articles included English abstracts thematically related to English teaching and learning during the COVID-19 pandemic.

3.2 Data Analysis

3.2.1 Analytical Tools

All analyses were conducted in RStudio (version 2024.12.1) using text mining packages. Key packages included *tidytext* for preprocessing and tokenization, *widyr* for co-occurrence and correlation analysis, *tidygraph* for network visualization, and *topicmodels* for Latent Dirichlet Allocation (LDA). The optimal number of topics was determined using the *ldatuning* package.

3.2.2 Text Preprocessing

Preprocessing is a critical step in text mining to ensure the accuracy and reliability of the findings. Raw textual data needs to be transformed into an appropriate form to address specific research problems (Romero and Ventura 2020). First, only nouns were extracted as key terms, since, as noted by Martin and Johnson (2015), they generally carry the main semantic content of a text and serve as key indicators of topics and entities. However, two adjectives, synchronous and asynchronous, were also retained due to their high frequency and conceptual significance during the pandemic, as they represent key instructional modes frequently discussed in online learning research. Next, stop words (e.g., research, study, purpose) were removed, abbreviations were standardized, and synonyms were unified. Compound terms (e.g., flipped learning, self-directed learning, self-directed, English-speaking) were

© 2026 KASELL All rights reserved

Table 1. Overview of Text Mining Studies in ELT

No.	Author (year)	Research Data (year range)	Topic	Text Mining Techniques	Analysis Tools
1	Chen et al. (2021)	1,295 articles on CALL (1995–2019)	Research trends, future directions in CALL	Structural Topic Modeling (STM) Bibliometric analysis Mann-Kendall trend test hierarchical clustering	Unspecified
2	CheshmehSohrabi and Mashhadi (2023)	2,162 documents and 1,903 articles (1973–2020)	Research trends	N-Gram analysis Clustering Bibliometric analysis (co-authorship network, citation analysis)	Excel, VOSviewer, RapidMiner software
3	Fujii et al. (2022)	647 Japanese EFL university students	Learner Satisfaction in Synchronous vs. Asynchronous Online EFL Courses	Co-occurrence Network Analysis	KH Coder 3.0
4	Hu et al. (2024)	423 WOS articles (2005–2024)	Changes in English education policy under the Belt and Road Initiative	LDA Topic Modeling, Word2Vec Cosine Similarity Analysis	Python (NLTK, Gensim, Word2Vec, Sankey diagram visualization)
5	Isti'anah and Wijanarka (2023)	444 Scopus-indexed articles (2020–2022)	Research trends in English teaching during COVID-19	Bibliometric analysis Co-authorship network analysis Keyword co-occurrence analysis	VOSViewer
6	Lee et al. (2023)	248 KCI-indexed articles (2010–2023)	Research Trend before & after Covid-19	Frequency Analysis Centrality Analysis Topic Modeling	Netminer 4.4.8v
7	Zhang and Hwang (2023)	20 Chinese EFL teachers	Teachers' identity reconstruction	Co-occurrence network Correspondence Analysis Collocation Statistics with Modal Verbs Jaccard Coefficient	KH Coder 3.0
8	Zhu (2025)	Educational texts, student feedback, teaching materials	Using text mining and topic modeling for curriculum enhancement in English teaching	IMT-HDPM (Intelligent Mined Text - Hierarchical Dirichlet Process Modelling) QTFM (Quantitative Term-Frequency Matrix)	Python 3.12.6v (NLP, TF-IDF, HDPM)

© 2026 KASELL All rights reserved

carefully preserved, as extracting only the noun component (e.g., learning) would fail to capture the contextual and pedagogical significance these expressions convey in ELT. Since such adjective–noun and noun–noun combinations function as core concepts in this field, it was necessary to classify them as distinct lexical units rather than separate them into single, generalized terms (Jo and Sim 2022, Park 2021b).

3.2.3 Analytical Procedures

This study conducted three layers of analysis: frequency analysis, semantic network analysis, and LDA topic modeling. These methods allowed for a comprehensive understanding of commonly used terms, their contextual associations, and the underlying thematic structures.

First, word frequency analysis was performed to identify the most frequently occurring terms in the dataset. Frequency analysis is widely used for identifying important terms in text summarization (Nenkova and McKeown 2012) and helps to identify which concepts the authors emphasized and considered important within the documents (Kim 2021).

Second, semantic network analysis was employed to capture the contextual associations among terms. This technique identifies key nodes and the relationships among them to understand the overall thematic structure of the dataset. To this end, two specific methods were used: co-occurrence analysis and the phi coefficient. Co-occurrence analysis assumes that the frequent co-appearance of two terms within a text indicates a meaningful thematic association between them (Milojević et al. 2011), whereas the phi coefficient measures binary correlation by comparing how often two words appear together or not at all, relative to how often they appear independently (Pesonen et al. 2020). Together, these methods helped identify term pairs with stronger contextual relevance. The semantic network visualizes these associations, highlighting the structural relationships among key terms and revealing how research themes are interconnected within the corpus.

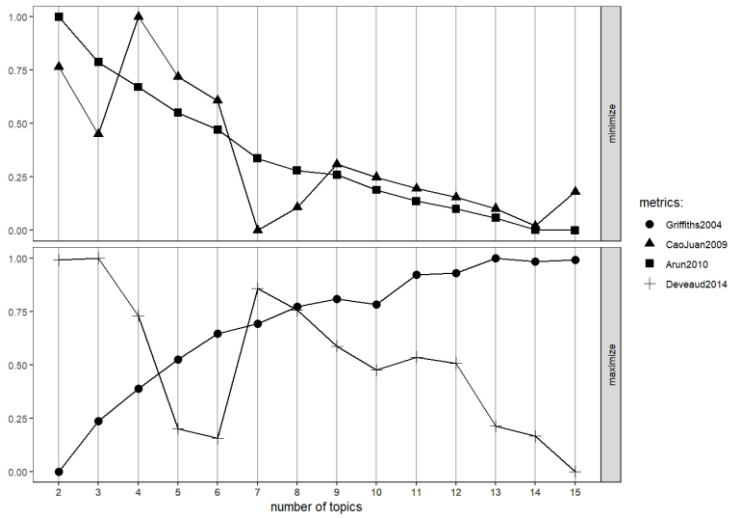

Figure 1. Topic Number Selection Using Four Evaluation Metrics

Finally, LDA topic modeling was used to identify latent thematic structures. It uncovers probabilistic topics in large-scale text corpora by clustering terms into topic groups based on their distribution across documents (DiMaggio et al. 2013). After running the LDA model, the optimal number of topics was determined using model evaluation metrics. To identify the optimal number of topics, Lim et al. (2021) used four widely recognized metrics:

© 2026 KASELL All rights reserved

Griffiths and Steyvers (2004) and Deveaud et al. (2014), for which higher scores indicate better performance, and Cao et al. (2009) and Arun et al. (2010), for which lower values indicate a better model fit. Based on these combined results, both a 7-topic and an 11-topic solution were considered (see Figure 1). The 7-topic model offered a more concise thematic overview, but the 11-topic model was ultimately selected for more detailed thematic interpretation.

Taken together, through these three layers of analysis, the study identified not only the most salient topics but also how they were interconnected, providing a holistic view of ELT research in Korea during the pandemic.

4. Results

4.1 Word Frequency Analysis

A total of 11,205 words were extracted from the dataset, and their frequency distribution is presented in Table 2. Two types of frequency were calculated: raw frequency refers to how often a word appears in the entire corpus, while document frequency indicates the number of documents in which a word appears. The three most frequent terms were "learner" ($n = 992$; 9.0 %), followed by "online" ($n = 812$; 7.3 %) and "learning" ($n = 533$; 4.8 %). In terms of document frequency, "learner" appeared in 98.4% of documents (91.6% for "online"; 76.8% for "learning"). These results indicate that research during the pandemic predominantly focused on learner-related issues, online delivery modes, and learning processes. A high document frequency for terms like "perception," "effect," and "level" further points to recurrent interests in learner perceptions, instructional effectiveness, and level-differentiated approaches.

Table 2. Top 20 Keywords by Raw and Document Frequency

Rank	Keyword	Raw Frequency (n=11,205)	Rank	Keyword	Document Frequency (n=190)
1	learner	992(9.00%)	1	learner	187(98.4%)
2	online	812(7.37%)	2	online	174(91.6%)
3	learning	533(4.83%)	3	learning	146(76.8%)
4	instructor	282(2.56%)	4	survey	139(73.2%)
5	survey	207(1.88%)	5	instructor	112(58.9%)
6	synchronous	189(1.71%)	6	course	76(40.0%)
7	group	174(1.58%)	6	teaching	76(40.0%)
8	course	160(1.45%)	8	perception	75(39.5%)
8	teaching	160(1.45%)	9	effect	73(38.4%)
10	satisfaction	155(1.41%)	10	synchronous	72(37.9%)
11	interaction	142(1.29%)	11	level	59(31.1%)
12	perception	134(1.22%)	11	satisfaction	59(31.1%)
13	offline	133(1.21%)	13	asynchronous	56(29.5%)
14	asynchronous	132(1.20%)	13	participant	56(29.5%)
15	lecture	129(1.17%)	15	group	55(28.9%)
16	level	124(1.12%)	16	offline	53(27.9%)
17	video	121(1.10%)	17	interaction	49(25.8%)
18	effect	117(1.06%)	18	environment	48(25.3%)
19	writing	101(0.92%)	19	video	47(24.7%)
20	participant	92(0.83%)	20	difference	45(23.7%)

© 2026 KASELL All rights reserved

4.2 Semantic Network Analysis

4.2.1 Co-occurrence Frequency Analysis

Table 3 presents the frequency of co-occurring word pairs across the documents. Among the top 20 pairs, "learner" appeared in nine, with the most common combinations aligning with the high-frequency terms identified earlier (e.g., "learner," "online," "learning"). For instance, the pair "learner–online" occurs in 91.1 % of documents, followed by "learner–learning" (75.8 %) and "learner–survey" (72.1 %). These results show that research explored learners' online learning experiences through survey methods.

Table 3. Co-occurrence Frequency of Word Pairs across the Documents (n = 190)

Rank	Co-occurring Word Pair	Number of Documents	Rank	Co-occurring Word Pair	Number of Documents
1	learner online	173(91.1%)	11	learner course	76(40.0%)
2	learner learning	144(75.8%)	12	learner teaching	75(39.5%)
3	learner survey	137(72.1%)	13	learner perception	74(38.9%)
4	learning online	135(71.1%)	14	learner effect	73(38.4%)
5	online survey	128(67.4%)	15	online course	72(37.9%)
6	learner instructor	110(57.9%)	16	learner synchronous	72(37.9%)
7	learning survey	108(56.8%)	17	online perception	71(37.4%)
8	online instructor	104(54.7%)	18	online teaching	70(36.8%)
9	learning instructor	88(46.3%)	19	online synchronous	68(35.8%)
10	instructor survey	86(45.3%)	20	online effect	67(35.3%)

To further examine the relationships among frequently co-occurring words, a semantic network graph was generated using co-occurring word pairs that appeared in more than 40 documents (see Figure 2). In this graph, each node represents a keyword, and each edge indicates a co-occurrence relationship between keywords. The node size and edge thickness reflect the frequency and strength of co-occurrence, respectively. Node colors indicate clusters of closely related terms. This visualization offers a clear overview of how key concepts, such as "learner," "online," "instructor," and "survey," are thematically interconnected in the documents.

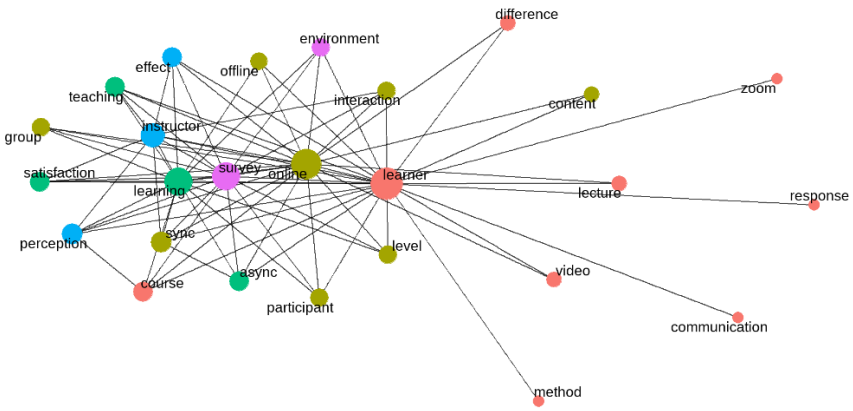

Figure 2. Semantic Network Graph of Co-occurring Word Pairs (Appeared in ≥ 40 Documents)

© 2026 KASELL All rights reserved

4.2.2 Correlation Analysis Using Phi Coefficient

The Phi coefficient ranges from –1 to 1, where values closer to 1 indicate a strong association between two words, values near 0 suggest a weak or no relationship, and values near –1 imply mutual exclusivity (Kim 2021). Table 4 presents the top 20 word pairs with the highest phi coefficients. Unlike frequency and co-occurrence analyses, the phi coefficient analysis indicates more tightly coupled conceptual associations. The top-ranking pair is "lecture–video," suggesting the prominence of video as a primary mode of lecture delivery. This is followed by "score–test," which may reflect the research interest in measuring learning outcomes, and "asynchronous–synchronous," indicating comparisons between synchronous and asynchronous formats. Notably, the pairs "beginner–level" and "beginner–proficiency" rank 4th and 5th, respectively, highlighting a notable research emphasis on novice learners.

Table 4. Top 20 Word Pairs by Phi Coefficient

Rank	Word Pair	Phi Coefficient	Rank	Phi Coefficient Word Pair	Phi Coefficient
1	lecture video	0.5716	11	method teaching	0.2903
2	score test	0.4504	12	level proficiency	0.2845
3	asynchronous synchronous	0.4469	13	synchronous zoom	0.2776
4	beginner level	0.4436	14	experience interview	0.2769
5	beginner proficiency	0.4141	15	group writing	0.2768
6	asynchronous video	0.3785	16	effect test	0.2683
7	achievement test	0.3530	17	learner online	0.2657
8	instructor interaction	0.3208	18	question response	0.2641
9	asynchronous lecture	0.3174	19	asynchronous type	0.2586
10	survey synchronous	0.3018	20	synchronous video	0.2562

4.3 LDA Topic Modeling Analysis

Table 5 presents the results of the LDA topic modeling, which extracted 11 major topics from the dataset. These topics were grouped into three broad thematic categories: learner-related, instructional design, and instructor-related themes. Six topics (55.2%) were categorized as learner-related, including self-directed learning, learner satisfaction, learning outcomes, speaking experiences in Zoom classes, learner achievement and proficiency, and anxiety and motivation. Four topics (35.2%) were associated with instructional design, covering course modalities and delivery formats (synchronous and asynchronous), the effectiveness of online reading instruction, and digital tools and learning environment design. One topic (9.5%) was classified as instructor-related, focusing on instructional methods and teacher communication.

First, in the learner-related category, the most frequently occurring topic focused on self-directed learning and learner engagement, which included keywords such as "self-directed," "satisfaction," "activity," and "motivation." This was followed by topics related to learner satisfaction with online learning and experiences using synchronous platforms such as Zoom. Other learner-related topics addressed learning outcomes (e.g., achievement, proficiency) and emotional factors (e.g., anxiety, motivation).

Second, within the instructional design category, distinct topics emerged related to course modalities, lesson formats, and digital tools. These topics included keywords such as "platform," "video," "blended-learning," and

© 2026 KASELL All rights reserved

"tool," suggesting that the studies in this domain focused primarily on course design and delivery modes.

Finally, the instructor-related category included only one topic, which focused on the instructor's teaching practices and communication, encompassing themes such as instructional methods, instructor–learner interaction, and instructors' use of English in instruction. This indicates that some studies explored pedagogical practices and teacher roles in online environments.

Overall, Table 5 shows that the majority of research during the pandemic was directed toward understanding learner experiences and the implementation of instructional strategies for digital learning environments.

Table 5. Topic Modeling Results: Major Themes, Keywords, and Distribution

Category	Topic Title	Top Keywords	Number of Articles	Proportion (%)
Learner-related	Self-directed Learning and Engagement	learning, learner, factor, self-directed, ability, self-efficacy	24	12.6
	Learner Satisfaction	learner, satisfaction, interaction, flipped-learning, model, design	18	9.5
	Speaking Experience in Zoom Classes	zoom, speaking, experience, response, question, classroom	18	9.5
	Level-based Writing Tasks and Feedback	group, learner, writing, activity, difference, feedback	17	8.9
	Proficiency and Achievement	level, learner, test, survey, proficiency, achievement	16	8.4
	Anxiety and Motivation Factors	learner, strategy, anxiety, learning, motivation, emotion	12	6.3
Subtotal			105	55.2
Instructional design	Course Modalities and Preferences	online, course, offline, perception, preference, blended-learning	22	11.6
	Effectiveness of Online Reading Instruction	online, effect, reading, comprehension, need, learner	16	8.4
	Synchronous and Asynchronous Formats	sync, async, lecture, video, survey, form	16	8.4
	Digital Tools and Learning Environment Design	content, environment, tool, perception, platform, technology	13	6.8
Subtotal			67	35.2
Instructor-related	Instructional Practices and Communication	instructor, teaching, communication, survey, method, English-speaking	18	9.5
Total			190	100

5. Discussion

5.1 Learner-Centered Focus and Attention to Novice Learners

In the frequency and co-occurrence analyses, the most frequently appearing keywords were "learner" and

© 2026 KASELL All rights reserved

"online." This pattern indicates a noticeable pattern of "learner-centered research", particularly in how learners engaged with and adapted to online environments during times of educational disruption. As shown by Park (2021a, 2021b), learner-centeredness and learner satisfaction were already expanding themes prior to the pandemic. However, the present study indicates that during the pandemic, the "learner" theme emerged as the central node of the entire research network. This resulted in a more densely connected network in which key instructional elements were mediated by the learners' digital adaptation.

This learner-centered tendency was also evident in the semantic network analysis, which identified four major thematic nodes: "learner," "online," "learning," and "survey." The "learner" cluster is closely linked to terms such as "course," "lecture," "video," "zoom," "method," "communication," and "response." These strong associations suggest a research focus on content delivery mechanisms and the role of synchronous tools in facilitating remote learning. Notably, the links to "communication" and "response" highlight the heightened importance of interaction in online settings, where communication and engagement became central elements of learning. These findings imply that studies primarily examined how learners interacted with different instructional formats and communication technologies in online learning environments.

Of particular significance, the phi coefficient analysis highlights associations related to "novice learners." Strong pairings such as "beginner–level" ($\varphi = 0.4436$) and "beginner–proficiency" ($\varphi = 0.4141$) point to an increased focus on how low-proficiency learners coped with the transition to online learning. While earlier studies identified learner proficiency as an important factor (Lee et al. 2023, Park 2021a), the present study identifies a marked emphasis. According to Park's (2021a) time-series analysis, the topic of "curriculum and level-differentiated programs" did not show an increasing trend ($p = .725$, ns). In contrast, the current findings indicate growing attention to low-proficiency learners. These results point to an increasing emphasis on providing targeted support for such learners in digitally mediated settings. For example, instructors in the post-pandemic era are encouraged to consider challenges that novice learners often face, such as emotional, motivational, and self-regulatory difficulties in online environments.

5.2 Thematic Patterns Identified Through Topic Modeling

Topic modeling identifies three primary areas of research focus during the pandemic: learner-related themes (55.2%), instructional design (35.2%), and instructor-related themes (9.5%). Prior to the pandemic, ELT research at the tertiary level had already begun shifting toward learner-centered and digitally integrated approaches. Ha (2017) noted a transition from curriculum development and assessment to learner-focused and ICT-based practices after 2000, and Park (2021a) observed increasing attention to learners' affective factors and learner-centered education after 2011.

This trajectory continued, as demonstrated in the present study, through a broader and more diverse exploration of learner experience and engagement. This exploration included self-directed learning, self-efficacy, learner satisfaction, learning experiences with Zoom, learner proficiency and achievement, learning strategies, anxiety, motivation, and emotion. In addition, several studies examined learners' perceptions of different instructional formats. With the expanded adoption of real-time platforms such as Zoom alongside asynchronous and on-demand formats, research described how synchronous and asynchronous modalities could be adapted for online classes. In these contexts, digital tools such as educational apps and platforms functioned not merely as delivery mechanisms but as means of promoting interaction and engagement. Therefore, instructors are encouraged to prioritize accessibility and the thoughtful integration of such tools, tailoring digital learning environments to meet diverse learner needs.

© 2026 KASELL All rights reserved

Regarding learners' affective factors, topics such as learner satisfaction (9.5%) and anxiety and motivation factors (6.3%) emerged as significant research themes (see Table 5), reflecting a deepening academic interest. This finding extends Park's (2021b) observation that affective domain terms, including motivation, attitude, perception, and learner satisfaction, markedly increased in frequency and centrality during the 2010s. These results suggest that while researchers were deeply engaged in adapting instructional delivery methods to the online environment, a significant pedagogical concern lay in addressing the shifting affective factors of learners.

In summary, research prior to the pandemic was already trending toward learner-centered instruction and digital integration. Pandemic-period studies further intensified this focus, with online learning methodologies and learner engagement in digital spaces emerging as prominent areas of investigation. This emphasis illustrates how the pandemic prompted researchers to explore new digital pedagogies, observe learner experiences and engagement, and consider the long-term implications of technology-driven education, reaffirming that the learner remains at the center of educational innovation.

5.3 Pedagogical Implications for English Language Teaching in the Generative AI Era

The pandemic underscored the extent to which access to devices and stable connectivity shaped learning opportunities (Atmojo and Nugroho 2020, Hijazi and AlNatour 2021) in a global context. As courses moved fully online, high-performance computers, stable internet connections, and multiple digital tools became prerequisites for participation, often exacerbating educational inequalities. Therefore, mobile learning emerged as an important alternative during this time because it required fewer resources and offered greater accessibility (Saikat et al. 2021). However, given Korea's advanced ICT infrastructure, research in Korea during the pandemic primarily focused on learners' experiences, engagement, and perceptions, as technological limitations themselves were not a pressing concern.

With the advent of generative AI, technology integration no longer depends primarily on hardware availability but on pedagogical choices and learner engagement (Rifqiyah et al. 2025). Because generative AI tools can be accessed easily on mobile devices, the pedagogical focus should be placed on how learners are instructed to approach and make effective use of these tools, rather than on issues of access. This underscores that learner experience and engagement, rather than equipment or infrastructure, must remain the central focus in technology-enhanced ELT.

Recent studies also point to its potential of generative AI to reinforce learner-centered practices, for example, by fostering autonomy and offering individualized pathways (Monika and Suganthan 2024, Ullah et al. 2025). With personalized feedback and adaptive learning pathways now made possible through generative AI-powered tools (Jamshed et al. 2024), learners are positioned to take greater control of their own learning processes. In this new paradigm, they can design and regulate their studies independently, pose questions, and refine their understanding through continuous feedback, even outside the classroom and without direct teacher supervision.

Consequently, future research should examine how learners are redefined as designers and active agents of their own learning, and how pedagogical frameworks can best support this transformation. Yet, this potential for autonomy also demands critical literacy, the ability to discern unreliable or hallucinatory outputs, to evaluate AI-generated feedback, and to make informed decisions in managing one's own learning. For low-proficiency learners, in particular, such autonomy must be scaffolded through explicit guidance to prevent cognitive overload and disengagement. Ultimately, future pedagogy should aim to cultivate learners who can design, regulate, and critically engage with their own learning in collaboration with AI, rather than depend passively on it.

© 2026 KASELL All rights reserved

6. Conclusion

In the context of generative AI transforming education, this study employs text mining to examine thematic patterns in tertiary-level ELT research in Korea during the pandemic. To this end, three main analytic methods were applied to KCI-indexed articles published between 2020 and June 2024.

First, frequency analysis showed that “learner” (98.4% of documents), “online” (91.6%), and “learning” (76.8%) dominated the corpus, underscoring the centrality of learners’ learning experience. Second, semantic network analysis identified “learner” as the central node, most strongly linked with “online” (91.1% of documents) and “learning” (75.8%), and further highlighted concerns for low-proficiency learners, as “beginner–level” and “beginner–proficiency” pairs ranked among the top five associations. Finally, LDA topic modeling produced 11 thematic clusters, with learner satisfaction, engagement, and autonomy emerging alongside instructional design issues such as delivery formats, digital tools, and blended learning.

Taken together, the results highlighted an increasing emphasis on learner-centered research, with topics such as self-directed learning, learner satisfaction, and engagement in online environments. Instructional design emerged as another major area, focusing on course delivery formats, digital tools, and blended learning models. These findings suggest that the pandemic not only accelerated the digital transformation of ELT but also emphasized the importance of learner experience and engagement, a perspective that remains critical in the era of AI-integrated ELT.

While these findings provide important insights, several methodological limitations should be noted. First, conducting a longitudinal comparison with research from the pre-COVID-19 period using weighted values, such as TF-IDF, would clarify temporal shifts more effectively. Second, although this study employed text mining techniques, a certain degree of researcher judgment was necessarily involved in both the selection and interpretation of topics. To enhance objectivity in future studies, incorporating expert review during these stages may be beneficial. Lastly, as this study relied solely on quantitative analysis and decontextualized data, it could not capture the lived experiences of educators and learners. Future research could incorporate qualitative methods, such as interviews or case studies, to allow for a more in-depth exploration of these perspectives.

Even after the crisis subsided, models such as blended learning have continued to gain attention as sustainable forms of instruction. Moreover, in line with broader developments in educational technology, ELT has kept pace with trends such as mobile-assisted learning and the integration of AI-driven tools (Wang and Kabilan 2024). In an era where technology continues to reshape education, a fundamental question remains: Who learns, and how? The pandemic reminded us that the success of education depends not on technology itself, but on the context in which it is applied. In this generative AI-mediated learning environment, learners can become active participants who design, regulate, and critically evaluate their own learning with the support of AI tools. This shift calls for further research to define learners’ evolving roles and to explore how pedagogy can effectively support them in exercising autonomy within generative AI-mediated learning environments. The pandemic demonstrated the impact of digital learning at scale, and current engagement with generative AI highlights the need to ensure that innovation remains grounded in learner-centered principles. Technology may change how we teach, but not for whom we teach.

© 2026 KASELL All rights reserved

References

Alghamdi, R. and K. Alfalqi. 2015. A survey of topic modeling in text mining. *International Journal of Advanced Computer Science and Applications* 6(1), 147-153.

Arun, R., V. Suresh, C. E. Veni Madhavan and M. N. Narasimha Murthy. 2010. On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010)*, 6118.

Atmojo, A. E. P. and A. Nugroho. 2020. EFL classes must go online! Teaching activities and challenges during COVID-19 pandemic in Indonesia. *Register Journal* 13(1), 49-76.

Blei, D. M., A. Y. Ng and M. I. Jordan. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research* 3, 993-1022.

Bozkurt, A. and R. C. Sharma. 2020. Emergency remote teaching in a time of global crisis due to Coronavirus pandemic. *Asian Journal of Distance Education* 15(1), i-vi.

Cao, J., T. Xia, J. Li, Y. Zhang and S. Tang. 2009. A density-based method for adaptive LDA model selection. *Neurocomputing* 72(7-9), 1775-1781.

Chen, X., D. Zou, H. Xie and F. Su. 2021. Twenty-five years of computer-assisted language learning: A topic modeling analysis. *Language Learning and Technology* 25(3), 151-185.

CheshmehSohrabi, M. and A. Mashhadi. 2023. Using data mining, text mining, and bibliometric techniques to the research trends and gaps in the field of language and linguistics. *Journal of Psycholinguistic Research* 52(2), 607-630.

Crawford, J., K. Butler-Henderson, J. Rudolph, B. Malkawi, M. Glowatz, R. Burton, P. A. Magni and S. Lam. 2020. COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Learning and Teaching* 3(1), 1-20.

Deveaud, R., E. SanJuan and P. Bellot. 2014. Accurate and effective latent concept modeling for ad hoc information retrieval. *Document Numérique* 17(1), 61-84.

DiMaggio, P., M. Nag and D. Blei. 2013. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics* 41(6), 570-606.

Erarslan, A. 2021. English language teaching and learning during Covid-19: A global perspective on the first year. *Journal of Educational Technology and Online Learning* 4(2), 349-367.

Feldman, R. and J. Sanger. 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge: Cambridge University Press.

Ferreira-Mello, R., M. André, A. Pinheiro, E. Costa and C. Romero. 2019. Text mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9(6), e1332.

Fujii, S., M. Mitsugi, K. Nakamura, Y. Ono, T. Yamagami, N. Takeuchi, H. Ishizuka and R. Kibler. 2022. A comparative study of learner satisfaction in synchronous and asynchronous online courses among Japanese EFL learners. *Journal of Pan-Pacific Association of Applied Linguistics* 26(2), 97-120.

Griffiths, T. L. and M. Steyvers. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America* 101(suppl_1), 5228-5235.

Ha, M. A. 2017. Analysis of research trends on college English education in Korea. *The Journal of Studies in Language* 33(3), 453-476.

Ha, M. A. 2022. A study on the dilemmas and challenges of flipped-learning-based college English education in the post COVID-19 era. *Journal of the Korea English Education Society* 21(3), 231-258.

© 2026 KASELL All rights reserved

Hijazi, D. and A. AlNatour. 2021. Online learning challenges affecting students of English in an EFL context during COVID-19 pandemic. *International Journal of Education and Practice* 9(2), 379-395.

Hodges, C., S. Moore, B. Lockee, T. Trust and A. Bond. 2020. The difference between emergency remote teaching and online learning. *EDUCAUSE Review.* Available online at https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Hu, H., F. Li and Z. Luo. 2024. The evolution of China's English education policy and challenges in higher education: Analysis based on LDA and Word2Vec. *Frontiers in Education* 9, 1385602.

Isti'anah, A. and H. Wijanarka. 2023. Teaching English during the pandemic: Bibliometric analysis. *LLT Journal: A Journal on Language and Language Teaching* 26(2), 650-664.

Jamshed, M., A. S. M. Manjur Ahmed, M. Sarfaraj, and W. U. Warda. 2024. The Impact of ChatGPT on English Language Learners' Writing Skills: An Assessment of AI Feedback on Mobile. *International Journal of Interactive Mobile Technologies* 18(19), 18-36.

Jo, K. H. and C. Y. Sim. 2022. A policy-centered analysis of the research trends in primary English education using topic modeling. *Primary English Education* 28(2), 5-30.

Khreisat, M. N. 2022. English language learning strategies during COVID-19 in the Middle East: A systematic review. *Arab World English Journal (AWEJ)* 13(3), 56-71.

Kim, S. H. and J. Y. Lim. 2013. The current state of college English education in Korea. *Modern English Education* 14(2), 263-290.

Kim, Y. W. 2021. *Do it! R for data analysis*. Seoul: EasysPublishing.

Korean Council for University Education. 2023. *Composing and Operating of Liberal Arts Education Curriculum in Korean Universities*. Report No. RR-2023-13-738. Seoul: Korean Council for University Education.

Law, L. 2024. Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open* 6, 100174.

Lee, E. H., H. S. Ahn and M. R. Park. 2023. Research trend analysis in college general English program before and after the COVID-19 pandemic: Utilizing topic modeling. *English Teaching* 78(3), 125-145.

Lim, J. H., Y. J. Oh and M. L. Ahn. 2021. An analysis of research trends on maker education using LDA-based topic modeling. *Journal of Korean Association for Educational Information and Media* 27(3), 1189-1219.

Lo, C. K., P. L. H. Yu, S. Xu, D. T. K. Ng and M. S.-Y. Jong. 2024. Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments* 11(1), 50.

Mahyoob, M. 2020. Challenges of e-learning during the COVID-19 pandemic experienced by EFL learners. *Arab World English Journal* 11(4), 351-362.

Martin, F. and M. Johnson. 2015. More efficient topic modelling through a noun only approach. In *Proceedings of the Australasian Language Technology Association Workshop 2015*, 111-115.

Martin, R. I., W. A. Din, A. A. Ab Aziz and S. Swanto. 2021. Higher education ESL curriculum transition post-COVID-19 era: A systematic review. *International Journal of Education, Psychology and Counselling* 6(43), 320-331.

Milojević, S., C. R. Sugimoto, E. Yan and Y. Ding. 2011. The cognitive structure of library and information science: Analysis of article title words. *Journal of the American Society for Information Science and Technology* 62(10), 1933-1953.

Mishra, L., T. Gupta and A. Shree. 2020. Online teaching-learning in higher education during lockdown period of Covid-19 pandemic. *International Journal of Educational Research Open* 1, 100012.

Mishra, S., S. Sahoo and S. Pandey. 2021. Research trends in online distance learning during the COVID-19

© 2026 KASELL All rights reserved

pandemic. *Distance Education* 42(4), 494-519.

Monika, M. and C. Suganthan. 2024. A study on analyzing the role of ChatGPT in English acquisition among ESL learners during English language classroom. *Bodhi International Journal of Research in Humanities, Arts and Science* 8(2), 75-84.

Moorhouse, B. L. and L. Kohnke. 2021. Responses of the English-language-teaching community to the COVID-19 pandemic. *RELC Journal* 52(3), 359-378.

Nenkova, A. and K. McKeown. 2012. A survey of text summarization techniques. In C. C. Aggarwal and C. Zhai, eds., *Mining Text Data*, 43-76. Boston: Springer.

Park, E. H. 2021a. Research trends in college English education in Korea–A topic analysis using LDA topic modeling. *Korean Journal of General Education* 15(5), 169-183.

Park, E. H. 2021b. Topic analysis in college English education in Korea using text network analysis. *Multimedia-Assisted Language Learning* 24(4), 214-239.

Pesonen, A. K., J. Lipsanen, R. Halonen, M. Elovainio, N. Sandman, J. M. Mäkelä, N. Antypa, A. Bédard and L. Kuula. 2020. Pandemic dreams: Network analysis of dream content during the COVID-19 lockdown. *Frontiers in Psychology* 11, 573961.

Ramalingam, S., M. M. Yunus and H. Hashim. 2022. Blended learning strategies for sustainable English as a second language education: A systematic review. *Sustainability* 14(13), 8051.

Rifqiyah, F., G. K. Kassymova and L. M. Harti. 2025. A bibliometric and LDA topic modeling analysis of artificial intelligence in English language learning. *Journal of Technological Pedagogy and Educational Development* 2(1), 55-68.

Romero, C. and S. Ventura. 2020. Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10(3), e1355.

Saikat, S., J. S. Dhillon, W. F. W. Ahmad and R. A. Jamaluddin. 2021. A systematic review of the benefits and challenges of mobile learning during the COVID-19 pandemic. *Education Sciences* 11(9), 459.

Soon, H. C. and A. A. Aziz. 2022. Teaching English online during the Covid-19 pandemic: A systematic literature review (2020-2022). *International Journal of Academic Research in Progressive Education and Development* 11(2), 678-699.

Ullah, Z., S. Ali, S. Iqbal and T. Ahmad. 2025. The role of Chat-GPT in learners' autonomy: Challenges and prospects for ESL learners. *Journal of Applied Linguistics and TESOL* 8(1), 1751-1762.

Wang, Y. and M. K. Kabilan. 2024. Integrating technology into English learning in higher education: A bibliometric analysis. *Cogent Education* 11(1), 2404201.

Zhang, L. and Y. Hwang. 2023. "Should I change myself or not?": Examining (re)constructed language teacher identity during the COVID-19 pandemic through text-mining. *Teaching and Teacher Education* 127, 104092.

Zhu, J. 2025. Text mining and topic modelling in English teaching: Extracting key themes and concepts for effective curriculum development. *Journal of Computational Methods in Sciences and Engineering* 25(2), 1210-1222.

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

© 2026 KASELL All rights reserved