



Research Trends in English Writing: A Semantic Network Analysis

Eunhee Park (Pukyong National 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: March 21, 2023
Revised: April 12, 2023
Accepted: May 17, 2023

Eunhee Park
Lecturer, College of Liberal Arts,
Pukyong National University
Tel: +82-51-629-6927
Email: jasminepark7@gmail.com

ABSTRACT

Park, Eunhee. 2023. Research trends in English writing: A semantic network analysis. *Korean Journal of English Language and Linguistics* 23, 376-396.

This study explores the shifting landscape of English writing research from 2001 to 2020, utilizing a large-scale, data-driven methodology. Data from 1495 articles published in Korea over these two decades was collected and preprocessed with the Biblio data collector, then analyzed with Netminer. The approach involved four stages of semantic analysis: frequency, centrality, network, and modularity analyses. The findings highlight dynamic shifts in research focus. While keywords like ‘test’, ‘college students’, ‘vocabulary’, and ‘level-based’ remained consistent, the 2010s saw emerging themes like ‘task’, ‘textbook analysis’, ‘corpus analysis,’ ‘peer feedback’, and ‘genre-based approach’. Centrality analysis showed that in the 2000s, alongside ‘sentence’, ‘questionnaire’ held a central position with multiple nodes linked to it. In the 2010s, ‘questionnaire’ persisted as a central theme but was joined by ‘relationship.’ Network maps generated with Spring 2D and PFnet depicted these evolving interconnections. In the second period, ‘feedback’ emerged as a central theme, yet directly connected to only two nodes: ‘error’ and ‘peer feedback.’ Modularity analysis identified six research groups in each period, with the ‘questionnaire group’ being most significant in the 2000s and the ‘peer feedback group’ gaining prominence in the 2010s. This research illuminates the evolving trends in English writing research, underscoring the potential of big data-driven approaches to uncover key insights and patterns.

KEYWORDS

social network analysis, SNA, semantic network analysis, research trend, English writing, L2 writing, centrality, community, modularity, Netminer

1. Introduction

English writing education has garnered substantial attention in language acquisition and pedagogy since the introduction of the Output Hypothesis by Swain (1985, 2000). In our increasingly globalized world, proficiency in English writing is essential for effective communication, academic success, and professional development. This has led to growing interest in exploring various aspects of English writing education, including instructional methods, assessment practices, feedback strategies, and the influence of writing on language development.

Research on English writing education in Korea has seen remarkable growth in both quantity and diversity. Scholars and educators are actively investigating different aspects of writing instruction, aiming to enhance students' writing skills. Park (2021) reported a significant increase in English writing research in Korea from 2001 to 2020, with a rise from 334 papers in the 2000s to 1161 papers in the 2010s. This surge in research output underscores the need for a broad analysis to scrutinize the changing patterns and shifts in the research landscape over this significant period.

This study addresses a critical gap by extensively analyzing English writing research trends in Korea from 2001 to 2020. It provides insights into long-term patterns, research focus evolution, and the emergence of new interest areas. The division of the overall period into two distinct sub-periods, 2001-2010 and 2011-2020, allows for a comparative analysis over time. This approach helps identify significant changes or developments between these two periods, including shifts in research topics, methodologies, and pedagogical approaches.

Moreover, this study represents a pioneering effort in applying semantic network analysis to English writing research trends in Korea. This methodology, which examines relationships and interconnections between concepts and ideas, transcends traditional researcher-oriented qualitative analyses. Semantic network analysis uncovers hidden patterns, connections, and central themes within the extensive literature on English writing education. This method aids in comprehensively understanding the research landscape in English writing education in Korea.

The findings from this study will enrich knowledge on English writing education by illustrating the changing patterns and research priorities in the field. By identifying key themes, influential nodes, and emerging research areas, this study lays a foundation for future research directions, curriculum development, and instructional practices in English writing education in Korea. Additionally, the use of semantic network analysis opens new avenues for research methodology and analysis in second language writing research. This novel approach can inspire future studies to explore the network dynamics of other language education domains, leading to further advancements in the field.

The subsequent sections present the methodology used in this study, provide a detailed analysis of the research data, discuss the results and implications, and conclude with suggestions for future research in English writing education in Korea.

2. Literature Review

2.1 Studies on Research Trends in English Writing

The potential of second language writing for language learning has been emphasized since the advent of the output hypothesis (Swain 1985). Cumming (1990) and Swain and Lapkin (1995) supported this hypothesis, showcasing the effects of writing practices on second language acquisition (SLA). With the increasing prominence of writing in SLA,

experimental research on English writing has notably expanded. Several studies, including Klein and Boscolo's (2016), have investigated trends and approaches in writing as a learning activity. They found that meta-analyses over the past decade confirm the reliable effects of writing on learning, with several variables acting as mediators and moderators. In recent years, the concept of writing to learn (WTL) has evolved (Manchón 2011), incorporating theories and research that merge social and psychological processes.

Riazi et al. (2018) reviewed 272 empirical research articles from the *Journal of Second Language Writing*'s first quarter-century of publication, spanning 1992 to 2016. Their analysis revealed that the predominant theoretical orientations were cognitive, social, socio-cognitive, genre, contrastive rhetoric, and critical theories, with writing instruction and feedback being the most common research foci. Building on the topic of feedback, Loncar et al. (2023) examined 293 empirical studies published from 2015 to 2019 on technology-mediated feedback in L2 writing contexts. Most of these studies incorporated both local and global feedback, with the prevailing research orientations being descriptive or comparative and primarily employing non-experimental designs.

Li and Storch (2017) ventured into a novel area within second language writing, specifically focusing on the era of computer-mediated communication (CMC). They accentuated aspects such as affordances, multimodality, and collaboration, emphasizing that multimodal composition and web-based collaborative writing are pivotal research topics in the CMC context. From another perspective, Fatimah and Masduqi (2017) examined the current trends in English as a Foreign Language (EFL) writing research in Indonesia. They discerned three significant themes: research on the writing process, research scrutinizing writing products, and research applying a genre-based approach. These diverse studies collectively enrich our understanding of the wide-ranging research trends in writing across various contexts.

In addition to international studies, domestic research has also examined trends in English writing within Korea (Chae 2012, Kang 2006, Lim 2020, Ma 2015, Park and Jang 2014). Kang (2006) analyzed 35 English writing education papers from the *English Education* journal, published between 1965 and 2005. His study spotlighted the field's emphasis on process-oriented feedback and evaluation, advocating for a broader array of research approaches. Expanding on Kang's work, Ma (2015) extended the analysis to 2015 and examined 127 papers. Her study identified seven focal areas: 'writing characteristics', 'classroom instruction', 'assessment', 'technology use', 'skill integration', 'learner traits', and 'others.' The researcher noted a disproportionate focus on 'classroom instruction', which accounted for nearly 42% of the research.

Park and Jang (2014) investigated 67 English writing education papers from the *English Language and Literature Teaching* journal, published from 1995 to 2013. They categorized the research into seven areas, including text analysis, writer & writing process, feedback, instructional design and teaching methods, writing assessment, materials analysis, and miscellaneous. Their analysis revealed a noticeable gap in research on the writing process, evaluation, and textbook analysis. In a similar vein, Lim (2020) reviewed 91 reading and writing-related articles from *Modern English Education*, published between 2000 and 2019. Lim's findings indicated a substantial focus on 'text analysis' and 'teacher and learner traits', while recommending an increased emphasis on 'assessment' and 'technology use.'

Chae (2012) analyzed 55 English writing articles to ascertain recent trends in English academic writing research. The researcher identified effective interventions such as 'teacher feedback,' 'self-regulatory learning,' 'peer feedback,' and 'technology use,' and recommended the inclusion of various measures and detailed prompts in future research.

Looking back on these studies, it's clear they offer valuable insights into the evolution of English writing research in Korea. The concentrated focus on papers published in specific academic journals, as evidenced by the works of Kang (2006), Ma (2015), Park and Jang (2014), and Lim (2020), provides an in-depth perspective on the development of the field within these contexts.

The current study, however, expands upon this focus, encompassing all English writing-related studies published

in Korea during the 2000s and 2010s, providing a broader perspective. Rather than adopting a deductive approach to classify research topics, it leverages the power of the data-driven tool, Netminer, to inductively identify and include all topics within each study. This innovative approach allows for a more inclusive representation of the studies, even those spanning multiple topics. Furthermore, this study goes beyond mere classification of topics, revealing the interrelationships and groupings among them. This not only enhances our understanding of the individual topics but also provides a more comprehensive view of the evolving landscape of English writing research in Korea over the past two decades.

2.2 Semantic Network Analysis

This study utilizes semantic network analysis to discern the interrelations of keywords within English writing research data. Essentially, semantic network analysis is the application of social network analysis methods to textual data, substituting the usual focus on connections between individuals or objects for a focus on associations between words. Much like social network analysis aims to pinpoint influential entities and their relationships, semantic network analysis seeks to identify key words and their interconnectedness within the text data. When multiple words occur within a defined proximity, they're considered connected. This connection is then analyzed and can be visually depicted through network visualization.

The history of network visualization started with Moreno's sociogram (Moreno 1934), which is an early form of social network analysis. In his hand-drawn sociogram, Moreno (1934) depicted relationships among students in a class using shapes and lines (Figure 1). Male students were represented by triangles, females by circles, and reciprocal relationships indicated by arrow directions. The figure reveals gender-based clusters, with interspersing students connecting the groups, and two isolated female students. This intuitive representation demonstrated that visualized data can provide more potent insights than lengthy verbal explanations. This principle is applicable to the semantic network analysis in our study.

Building on this principle of visualizing relationships, the concept of centrality emerges as a crucial index in understanding networks. Centrality identifies nodes which are similar to the students in Moreno's sociogram, and these nodes hold critical positions within the network (Zhang and Luo, 2017). Nodes with high centrality are located near the center of the network, indicating their power and influence over other nodes. Freeman (1977, 1978/79) introduced degree centrality, closeness centrality, and betweenness centrality as commonly used measures in social network analysis. Degree centrality represents the number of connections a node has and higher degree centrality implies greater influence. Closeness centrality measures how central a node is located and its proximity to other nodes. The more central a node is, the more significant its role within the community. Betweenness centrality assesses the extent to which a node acts as a mediator in the community. While these centrality measures have distinct conceptual differences, Zhang and Luo (2017) observed that they often demonstrate similar results.

In the 1970s, centrality analysis emerged, followed by the development of a method for identifying communities in the 2000s (Newman 2006a, 2006b). The modularity technique has become widely used in large-scale network analysis for identifying groups even when clear criteria for defining the groups are absent. Figure 2, presented by Medaglia and Bassett (2017), illustrates an example of network analysis using modules. It reveals the division of the network into three modules interconnected by nodes that bridge each module. Medaglia and Bassett (2017) coined the term 'connector hub' for nodes connecting different modules and 'provincial hub' for nodes playing important roles within a module. This research utilizes both degree centrality and modularity analysis, aiming to incorporate these two concepts into the discussion.

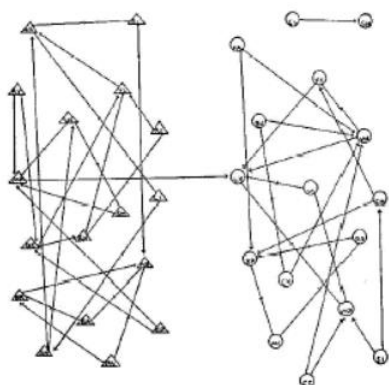


Figure 1. Friendship Choices Among Fourth Graders (Moreno, 1934, p. 38)

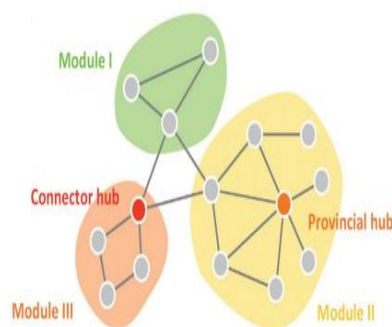


Figure 2. Types of Network Hubs (Medaglia and Bassett, 2017, p. 18.)

Semantic network analysis was applied to English education research data in a study conducted by Hwang and Kim (2019) and Shin and Kim (2020). Hwang and Kim (2019) examined domestic research trends in North Korean English education using keyword network analysis. They analyzed 83 papers published between 1993 and April 2019, identifying main topics for different periods: the North Korean English education situation (1990-1999), English textbook analysis (2000-2009), and studies on North Korean defector learners along with English textbook analysis (2010-2019). Shin and Kim (2020) conducted a keyword network analysis on 814 papers published in the journal, *English Education Research* over a 20-year period from 2000 to 2019. Their findings revealed a shift in research focus from the internal characteristics of learners and instructors (2010-2019) towards improving the learning environment and exploring the four language areas (2000-2009).

In the realm of English education, previous studies have delved into the application of network analysis. However, the current research sets itself apart by employing semantic network analysis and embarking on a novel path of inquiry. This study not only encompasses a significantly larger and more expansive dataset, covering all papers related to English writing from 2001 to 2020, but it also explores a distinct research domain. By analyzing the semantic connections within the data, this study aims to shed light on previously unexplored aspects and seeks answers to the following intriguing questions:

1. Which high-frequency words are prevalent in English writing research during the 2000s and 2010s?
2. Which words hold central positions and strong connections to other words within each period?
3. How does the network map of keywords appear and differ across periods?
4. How do communities form within the network, and are there periodical differences?

Through addressing these research questions, the study endeavors to reveal significant insights into the landscape of English writing research and contribute to the advancement of the field.

3. Research Method

3.1 Research Procedure

As shown in Figure 3, this study was conducted through three stages: data collection, data preprocessing, and network analysis.

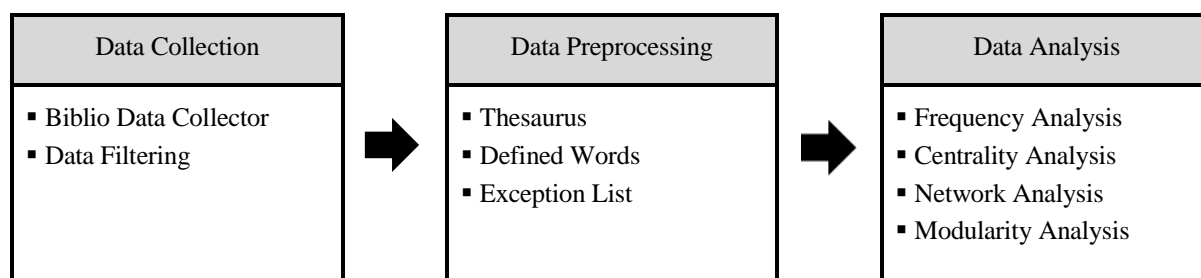


Figure 3. Research Framework

3.2 Data Collection

In this study, Netminer 4.4 and its extension program, Biblio Data Collector, were utilized for data collection and data analysis. The data collection targeted the KCI (Korean academic journal citation index) category. A specific search strategy was employed using the keywords ‘writing’, ‘write’, ‘writer’, ‘written’, and ‘composition’ in combination with ‘English’, ‘Second language’, and ‘L2.’ These keywords were connected using the ‘AND’ connector to refine the search. The initial search yielded 2,649 papers. Following this, the bibliographic data of the papers were scrutinized. Any papers that were deemed irrelevant to the research objectives were excluded by utilizing the Query function of Biblio Data Collector. This systematic filtering process resulted in a final dataset of 1,495 papers. For the purpose of this study, the English abstracts (Abstract_EN) of the collected papers were specifically targeted as the primary data source.

3.3 Data Preprocessing

Preprocessing, a crucial step for social network analysis, was undertaken in Netminer using three dictionaries: Thesaurus, Defined Words, and Exception List. The Thesaurus unified acronyms and similar expressions under one representative word to ensure consistency in analysis, for example, treating ‘corrective feedback’ and its acronym ‘CF’ as the same term. The Defined Words dictionary was used to identify phrases that should be recognized as single analytical units, such as ‘peer feedback’ and ‘high school.’ The Exception List excluded words that did not contribute to the study’s specific focus, like ‘analysis.’ For this study, 298 Thesaurus entries, 1131 Defined Words, and 210 Exception List words were generated. Table 1 provides examples of these dictionaries. Consistent with prior studies (Kwon 2021, Won and Kim 2021), the analysis targeted noun morphemes.

Table 1. Preprocessing Categories, Examples, and Number of Entries

Category	Examples	Number of Entries
Thesaurus	L1, first language; level-differentiated, level-based	298
Defined Words	textbook analysis; high school; written corrective feedback	1131
Exception List	analysis; data; English; importance; Korea; language, purpose	210

3.4 Data Analysis

3.4.1 Frequency Analysis

To illuminate prevalent topics, themes, and concepts in English writing research literature during the 2000s and 2010s, frequency analysis was employed. This technique offers a quantitative assessment of recurring ideas by measuring word prominence within each period. For the purpose of research question 1, frequency was computed in a specific way: instead of counting the total occurrences of a word throughout the dataset (simple frequency), the count was based on the number of papers in which a word appeared, regardless of how many times it was repeated within a single paper. For instance, if ‘corpus analysis’ was mentioned thrice in one paper’s abstract, it was counted as one occurrence. This approach, implementable through the Netminer program’s Analyze function, avoids potential distortion that could result from overemphasizing repeatedly used words within individual papers.

3.4.2. Centrality Analysis

In response to research question 2, a centrality analysis of 40 select words from each period was conducted subsequent to the frequency analysis. This examination aimed to pinpoint influential nodes, which are words with substantial connectedness and significance, in the semantic network of English writing research. Utilizing degree centrality, a widely used index denoting the number of nodes connected to a specific node (Kwon 2021, Zhang and Luo 2017, Kim and Kim 2019), the analysis identified key concepts shaping the research landscape. Through this approach, we gained a deeper understanding of each node’s centrality and influence within the network.

3.4.3. Network Analysis

In order to answer research question 3 and further illustrate the relationships and connections identified in the centrality analysis, a network map was created using PFnet and Spring 2D techniques. PFnet, a method that condenses the network while maintaining the shortest path between nodes, and Spring 2D, a commonly used graph algorithm akin to subway maps, were utilized for effective visualization. The chosen Spring 2D layout was based on Kamada and Kawai’s (1989) algorithm analysis method. These visualizations provided a comprehensive view of the interdependencies and associations among selected words, revealing key nodes and underlying structures in English writing research across the periods. By mapping the network, the study unveiled significant relationships and dynamics within the research discourse.

3.4.4. Modularity Analysis

To address research question 4 and discern the community structure within the network, modularity analysis was deployed. This method aims to identify distinct communities or clusters of words that exhibit similar characteristics, thereby clarifying the interconnections among various research topics and their contributions to the overall research landscape. In this context, a ‘community’ represents a subset of nodes with more internal than external interactions, while ‘modularity’ refers to the extent to which the network can be divided into these distinct communities (Fortunato and Barthélemy 2007, Girvan and Newman 2002, Jun and Han 2013, Newman 2018, Newman and Girvan 2004). A network with high modularity has well-separated communities, whereas one with low modularity exhibits greater interconnectedness. The study also integrated the concepts of ‘connector node’ and ‘provincial node’ proposed by Medaglia and Bassett (2017), providing insights into distinct research directions within each period.

Overall, the four analyses are interrelated and form a logical sequence of exploration. Starting with frequency analysis, prominent words were identified, followed by centrality analysis to determine their central power. The network analysis visualized the relationships between these words, and the modularity analysis revealed the grouping or clustering of words into distinct communities. Together, these analyses provide a comprehensive understanding of the research trends, influential concepts, and the overall structure of English writing research during the two periods under investigation.

4. Results and Discussion

4.1 Frequency analysis

Following preprocessing, a total of 5938 words were gathered. A frequency analysis was then performed to organize these words by their occurrence. As previously stated, ‘frequency’ is defined as the incidence of words across various papers. Table 2 showcases the 40 most frequently occurring words, along with their frequency counts during two periods: 2001-2010 (period 1) and 2011-2020 (period 2).

Table 2. High-Frequency Words per Period

2001-2010			2011-2020		
Rank	Keywords	Frequency Count	Rank	Keywords	Frequency Count
1	<i>difference</i>	82	1	test	92
2	<i>vocabulary</i>	73	2	college students	91
3	<i>questionnaire</i>	73	3	<i>difference</i>	90
4	test	67	4	<i>vocabulary</i>	89
5	<i>level-based</i>	67	5	feedback	82
6	<i>classroom</i>	62	6	<i>questionnaire</i>	80
7	college students	61	7	<i>level-based</i>	76
8	<i>sentence</i>	57	8	text	74
9	text	53	9	error	66
10	<i>interview</i>	49	10	essay	59

11	<i>process</i>	48	11	structure	52
12	university	45	12	term	50
13	error	45	13	task	50
14	term	44	14	score	50
15	score	44	15	<i>sentence</i>	47
16	essay	43	16	knowledge	46
17	development	43	17	development	46
18	feedback	40	18	university	43
19	<i>subject(-)</i>	38	19	assessment	42
20	<i>content(-)</i>	38	20	<i>interview</i>	41
21	<i>elementary school(-)</i>	37	21	proficiency	40
22	structure	36	22	program	37
23	knowledge	36	23	<i>process</i>	37
24	<i>attitude</i>	36	24	accuracy(+)	37
25	article	36	25	strategy	36
26	<i>writing skills</i>	35	26	corpus analysis(+)	36
27	strategy	35	27	<i>classroom</i>	36
28	proficiency	34	28	week(+)	35
29	<i>journal writing(-)</i>	33	29	quality(+)	35
30	grammar instruction	32	30	grammar instruction	35
31	<i>L1(-)</i>	31	31	high school students(+)	34
32	<i>relationship</i>	30	32	peer feedback(+)	33
33	<i>form(-)</i>	30	33	genre-based approach(+)	33
34	task	29	34	<i>article</i>	33
35	program	29	35	<i>writing skills</i>	32
36	<i>pattern(-)</i>	29	36	textbook analysis(+)	32
37	<i>topic(-)</i>	28	37	perception(+)	32
38	need(-)	28	38	corrective feedback(+)	32
39	assessment	28	39	<i>relationship</i>	31
40	<i>interests(-)</i>	27	40	<i>attitude</i>	31

A significant observation from the frequency analysis was that six words, namely ‘test’, ‘college students’, ‘difference’, ‘vocabulary’, ‘questionnaire’, and ‘level-based’, consistently appeared within the top seven in both periods 1 and 2. Of these, ‘test’, ‘difference’, ‘questionnaire’, and ‘level-based’ pertain to research design, while ‘college students’ corresponds to the study’s subject matter. This suggests that most writing research focused on college students, with experimental designs commonly used to examine group differences or effects of specific interventions. ‘Vocabulary’ stands out as the sole word associated with the research purpose, implying a significant number of studies in writing research are centered around vocabulary.

Upon comparison, several trends between the two periods emerge. Firstly, the frequency of words like ‘feedback’ (18th to 5th), ‘peer feedback’ (58th to 32nd), and ‘corrective feedback’ (257th to 38th) increased markedly in period 2. Similarly, feedback-related terms such as ‘error’ (13th to 9th) and ‘accuracy’ (55th to 24th) followed this upward trajectory, indicating an enhanced focus on the topic of feedback. Secondly, emerging domains in writing research such as ‘task’ (34th to 13th), ‘textbook analysis’ (45th to 23rd), ‘corpus analysis’ (90th to 26th), and ‘genre-based

achieved the highest centrality. This indicates that frequency and centrality don't always align, suggesting that centrality analysis can offer a distinct perspective on a network. For example, 'questionnaire' and 'sentence' appear to exert more influence than their occurrence frequency would imply. The terms 'classroom', 'college students', 'difference', and 'vocabulary' followed closely, each with a centrality index of 0.103. On the other hand, terms like 'L1', 'attitude', 'content', 'feedback', 'form', and 'journal writing' showed relatively low centrality.

In Figure 5, 'questionnaire' and 'relationship' exhibit the highest centrality for Period 2, each with a centrality index of 0.128. 'Questionnaire', despite ranking 6th in the frequency analysis for Period 2, occupied the top position in centrality analysis, indicating its prevalent use as a research method across both periods. Interestingly, 'relationship', ranking 39th in the frequency analysis for Period 2, emerged as one of the words with the highest degree of centrality in this period's analysis. This trend suggests an increase in experimental studies exploring the relationship between writing effects and mediating variables in Period 2 compared to Period 1. The terms 'error' and 'peer feedback' followed 'questionnaire' and 'relationship' in centrality during Period 2, each with a centrality index of 0.103.

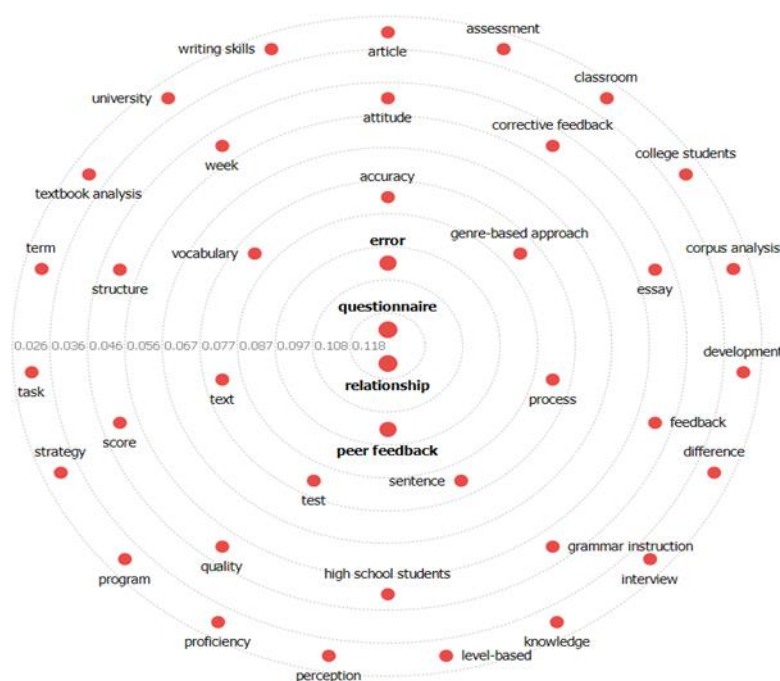


Figure 5. Centrality Analysis: 2011-2020

Table 3 presents 11 words that demonstrate high centrality during Periods 1 and 2, along with their respective centrality indices. In this table, words that share the same centrality index are assigned the same rank.

Table 3. Centrality Analysis Comparison: 2001-2020 vs. 2011-2020

2001-2010			2011-2020		
Rank	Words	Centrality	Rank	Words	centrality
1	questionnaire	0.128	1	questionnaire	0.128
1	sentence	0.128	1	relationship(+)	0.128
3	college students(-)	0.103	3	peer feedback(+)	0.103
3	classroom(-)	0.103	3	error(+)	0.103
3	vocabulary	0.103	5	accuracy(+)	0.077
3	difference(-)	0.103	5	vocabulary	0.077
7	assessment(-)	0.077	5	text	0.077
7	text	0.077	5	test	0.077
7	score(-)	0.077	5	sentence	0.077
7	interview(-)	0.077	5	process(+)	0.077
7	test	0.077	5	genre-based approach(+)	0.077

Words exhibiting high centrality in both periods include ‘sentence’, ‘vocabulary’, ‘text’, ‘questionnaire’, and ‘test.’ As previously noted in the frequency analysis, these terms were also frequently used across both periods, with the exception of ‘college students’, ‘difference’, and ‘level-based’, which were less influential. The prominence of ‘vocabulary’, ‘sentence’, and ‘text’ is logical given that writing inherently involves composing sentences with a variety of vocabulary. The commonality of ‘test’ and ‘questionnaire’ underscores the frequent use of these methodologies in English writing research design.

Interestingly, ‘grammar instruction’ maintained the same rank (12th) and index (0.051) in the centrality analysis across both periods, matching its consistent rank (30th) in the frequency analysis. This persistence suggests that the role and frequency of grammar instruction in writing research remain stable over time.

Several words displayed period-specific changes in centrality index. Terms with a lower centrality in Period 2 compared to Period 1 included ‘sentence’, ‘college students’, ‘classroom’, ‘vocabulary’, ‘difference’, ‘assessment’, ‘score’, and ‘interview.’ Conversely, words that saw increased centrality in Period 2 were ‘relationship’, ‘peer feedback’, ‘error’, ‘accuracy’, ‘text’, ‘test’, ‘process’, and ‘genre-based approach.’ These shifts signify important changes in research topics over time.

During the 2000s, research primarily revolved around foundational topics related to English education, such as vocabulary learning and classroom-based studies. However, the focus shifted in the 2010s to more specific writing-related topics, including the writing process, peer feedback, and genre-based writing. This transition suggests that writing research became more specialized and nuanced in the 2010s.

Moreover, it's interesting to contrast the centrality of ‘peer feedback’ and ‘feedback’ in Period 2. In the frequency analysis, ‘feedback’ claimed the 5th spot while ‘peer feedback’ was ranked 32nd. However, in the centrality analysis, ‘peer feedback’ jumped to 3rd place (0.103), surpassing ‘feedback’ which dropped to 12th (0.051). Despite the lesser frequency of ‘peer feedback’, it was more central within the network, indicating its higher influence and popularity. This reinforces the idea that high frequency doesn't necessarily equate to strong centrality.

Not only do the comparisons within the same period reveal differences, but the comparison between the frequency analysis and centrality analysis also yields varying results. For instance, in Period 1, words such as ‘test’, ‘vocabulary’, and ‘questionnaire’ appeared in both frequency and centrality rankings, while other words differed between the two. In Period 2, only a limited number of words, including ‘relationship’, ‘peer feedback’, and ‘accuracy’, were listed in

1, holds the most influence within the network. A key aspect to observe in the network analysis is the connections between nodes. In the case of ‘questionnaire’, it is connected to five nodes: ‘interview’, ‘attitude’, ‘interest’, ‘test’, and ‘writing skills’. This suggests that during Period 1, numerous studies incorporated questionnaires in conjunction with interviews, exploring learners’ attitudes or interests, as well as studies that employed questionnaires before or after tests.

Similarly, ‘sentence’, which possesses the highest centrality in Period 1, much like ‘questionnaire’, is also connected to five nodes: ‘structure’, ‘strategy’, ‘level-based’, ‘error’, and ‘vocabulary.’ This suggests that research endeavors at the sentence level have encompassed aspects such as text structure, writing strategies, error analysis, and vocabulary acquisition. Despite sharing the same centrality, ‘questionnaire’ and ‘sentence’ occupy different positions within the network. While ‘questionnaire’ resides in the central region of the map, ‘sentence’ is positioned in the southeast, somewhat removed from the center. To provide an analogy, both ‘questionnaire’ and ‘sentence’ have numerous associations, but the group led by ‘questionnaire’ holds more prominence and influence within the network.

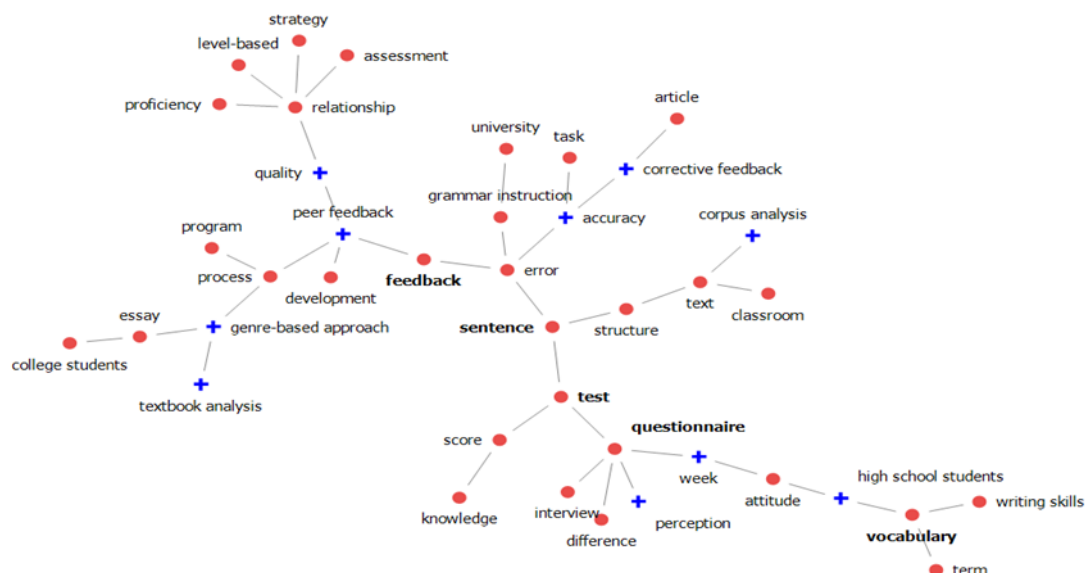


Figure 7. Network Analysis: 2011-2020

Figure 7 depicts the network map for Period 2. Similar to Figure 6, nodes in Figure 7 are denoted by red circles or blue plus marks. Red circles represent words that appear in both Period 1 and Period 2, while blue plus marks indicate words that emerged in Period 2. Specifically, the following words are newly introduced in Period 2: ‘quality’, ‘peer feedback’, ‘accuracy’, ‘corrective feedback’, ‘corpus analysis’, ‘week’, ‘high school students’, ‘perception’, ‘genre-based approach’, and ‘textbook analysis.’ These additions reflect the evolving landscape of research topics and areas of focus during Period 2.

Figure 7 clearly demonstrates that ‘feedback’ holds the most central position within the network. Interestingly,

'feedback' transitions from being a peripheral node located at the left end in the 2000s (Figure 6) to occupying a central position in the 2010s. Notably, 'feedback' maintains a connection with the 'error' node in both periods. This suggests that the primary focus of feedback, throughout both the 2000s and 2010s, has been on error correction. Despite the different positions occupied by 'feedback' in each period, its consistent connection with 'error' underscores the continued emphasis on error correction across both time periods.

The dynamics between 'feedback', 'peer feedback', and 'corrective feedback' are intriguing. In Figure 7, 'feedback' and 'peer feedback' share a direct connection, while 'corrective feedback' is positioned further apart with three or four nodes in between. Remarkably, 'peer feedback' occupies a central position within the map and is connected to four nodes, whereas 'corrective feedback' is located in the northeast section of the map and linked to two nodes. This suggests that 'peer feedback' exhibits higher subject connectivity, while 'corrective feedback' demonstrates greater subject independence.

Furthermore, despite 'questionnaire' and 'relationship' being the words with the highest centrality in Period 2, they yield the central position to 'feedback' and are situated towards the periphery of the map.

Research question 3 aimed to visualize the network connections between words. In Period 1, the central node of the network map was 'questionnaire', which exhibited links to 'interview', 'attitude', 'interest', 'test', and 'writing skills.' In Period 2, the central node shifted to 'feedback.' Intriguingly, 'feedback' was connected solely to 'error' in Period 1, while in Period 2, it was linked to two nodes, 'error' and 'peer feedback.' The consistent connection between 'feedback' and 'error' across both periods underscores the primary focus on error correction in writing research.

4.4 Modularity Analysis

Modularity analysis was employed to examine the extent to which the network could be divided into distinct communities. The analysis of Period 1 networks resulted in the identification of six groups, as depicted in Figure 8. Each group is represented by a blue pentagon, indicating the presence of representative nodes within the respective groups. In this context, the term 'representative node' refers to nodes that fulfill the roles of both provincial nodes and connector nodes, following the framework proposed by Medaglia and Bassett (2017).

The representative nodes for each group in Period 1 were 'sentence', 'vocabulary', 'questionnaire', 'classroom', 'test', and 'text.' Notably, four out of the six representative nodes served as both provincial nodes and connector nodes: 'sentence', 'questionnaire', 'classroom', and 'test'. As illustrated in Figure 8, each module includes at least one node assuming a provincial role, as well as a connecting node ensuring intermodule connectivity.

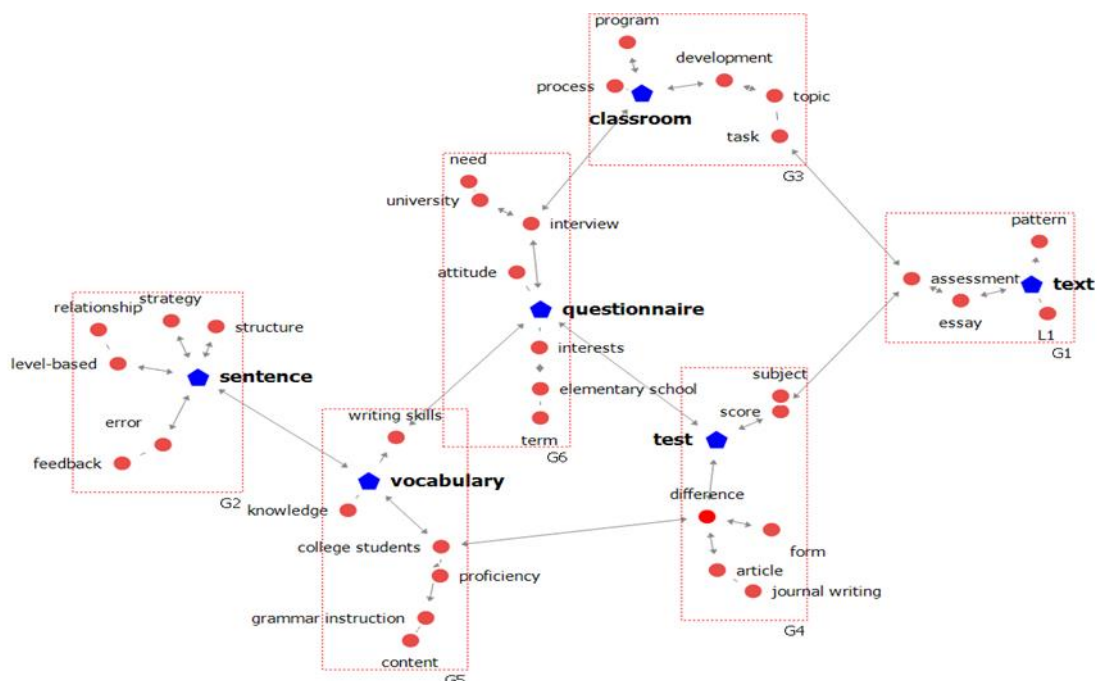


Figure 8. Modularity Analysis: 2001-2010

Table 4 presents the properties of each group identified through modularity analysis. The ‘Members’ column lists the nodes belonging to each group, while the ‘Size’ column indicates the number of members. ‘Density’ refers to the internal density of a group, ranging from 0 to 1, with higher values indicating increased density. The ‘Cohesion Index’ is derived by comparing the link density within a group to the link density outside the group. An index of 1 signifies equal link densities within and outside the group, while a value greater than 1 suggests higher link density within the group.

The group with the largest size was G6, consisting of nodes such as ‘questionnaire’, ‘attitude’, ‘interests’, ‘interview’, ‘need’, ‘university’, ‘elementary school’, and ‘term.’ G6 is associated with the research method, research subject, and the affective domain of the learners. In terms of density, G1 exhibited the highest value and included ‘Text’, ‘L1’, ‘assessment’, ‘essay’, and ‘pattern’. This group focuses on the influence of L1 on English writing and the analysis and evaluation of text units. Additionally, G2 displayed the highest cohesion index, with members such as ‘sentence’, ‘error’, ‘feedback’, ‘level-based’, ‘relationship’, ‘strategy’, and ‘structure.’ G2 represents the densest group, emphasizing linguistic aspects with stronger internal connections compared to other groups.

Table 4. Group Properties: 2001-2010

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Size	5	7	6	7	7	8
Density	0.116	0.083	0.083	0.081	0.084	0.077
Cohesion	17.5	33	17	11	11	10.667
Members	Text	Sentence	Classroom	Test	Vocabulary	Questionnaire
	L1	error	development	score	college students	attitude
	assessment	feedback	process	subject	content	elementary school
	essay	level-based	program	difference	grammar instruction	interests
	pattern	relationship	task	form	knowledge	interview
		strategy	topic	article	proficiency	need
		structure		journal writing	writing skills	term
						university

Like Period 1, the network in Period 2 was also subjected to modularity analysis, resulting in the identification of six groups, as depicted in Figure 9.

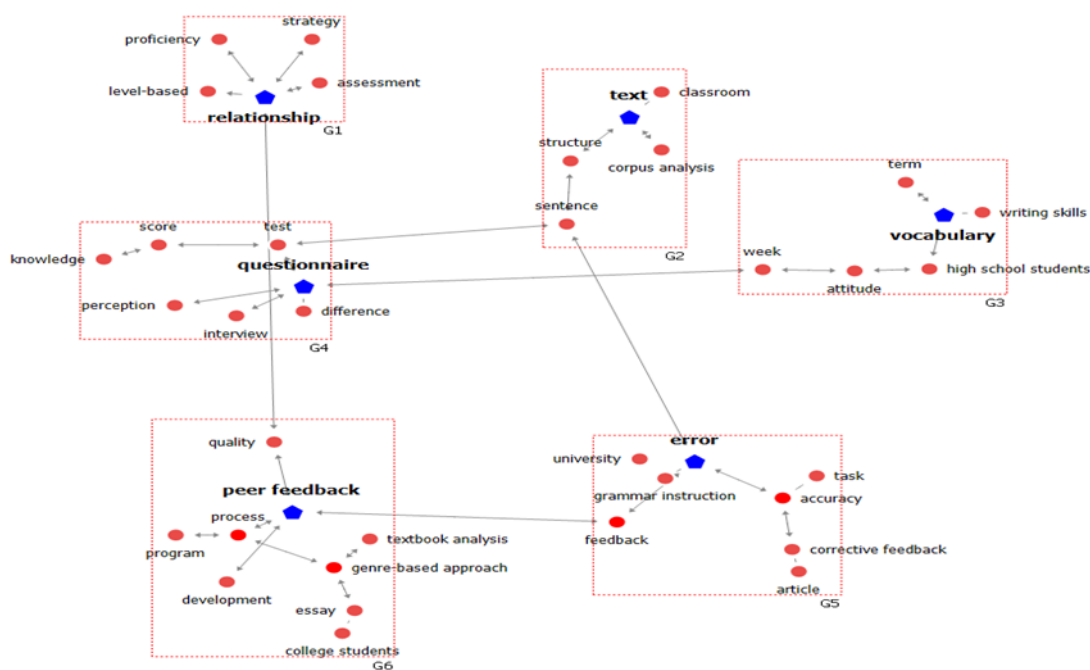


Figure 9. Modularity Analysis: 2011-2020

In a similar manner to Period 1, representative nodes were selected for each group in Period 2 using the same

methodology. The representative nodes for Period 2 were ‘relationship’, ‘text’, ‘vocabulary’, ‘questionnaire’, ‘error’, and ‘peer feedback.’ Remarkably, three out of the six representative nodes were also representative nodes in Period 1: ‘text’, ‘vocabulary’, and ‘questionnaire.’ However, ‘relationship’, ‘error’, and ‘peer feedback’ emerged as new groups in Period 2. Figure 9 illustrates that, like in Period 1, most of the representative nodes serve as connector hubs and/or provincial hubs within the network.

Table 5 presents the properties and members of each group in Period 2. The largest group, G6, consisted of nine members: ‘peer feedback’, ‘college students’, ‘development’, ‘essay’, ‘genre-based approach’, ‘process’, ‘program’, ‘quality’, and ‘textbook analysis.’ Similar to G6 in Period 1, G6 in Period 2 exhibited the lowest density and cohesion index compared to other groups.

On the other hand, both G1 and G3 had cohesion indices more than twice as high as those of other groups. This indicates that the density within these groups was significantly higher than that outside the groups. Consequently, it can be inferred that the relationship between learner level and evaluation (G1) and vocabulary instruction in a writing class (G3) exhibit higher subject independence compared to other groups.

Table 5. Group Properties: 2011-2020

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Size	5	5	6	7	8	9
Density	0.045	0.043	0.038	0.041	0.033	0.027
Cohesion	35	17.5	34	16.5	16	15.5
Members	relationship assessment level-based proficiency strategy	text classroom corpus analysis sentence structure	vocabulary attitude high school students term week writing skills	questionnaire difference interview knowledge perception score test	error accuracy article corrective feedback grammar instruction task university	peer feedback college students development essay genre-based approach process program quality textbook analysis

Research question 4 aimed to examine how words in the network were organized using modularity analysis. Both in Period 1 and Period 2, six distinct groups were identified within the network. Each group consisted of one or two words that were connected to words from other groups, resulting in a fully connected network. The ‘text group’, ‘questionnaire group’, and ‘vocabulary group’ were consistently present in both periods. In addition to these common groups, Period 1 featured the ‘sentence group’, ‘classroom group’, and ‘test group’, while Period 2 included the ‘relationship group’, ‘error group’, and ‘peer feedback group.’ Among these groups, the ‘questionnaire group’ was the largest in Period 1, whereas the peer feedback group emerged as the largest in Period 2.

5. Conclusion

This study has undertaken a comprehensive analysis of English writing research in Korea, examining the trends and developments over two distinct periods, the 2000s and the 2010s. Research papers were analyzed, and semantic network analysis was employed, revealing key shifts in the focus of English writing education and research in Korea between the two decades. The study began with a frequency analysis to identify the most common words. These words were then subjected to a centrality analysis to assess their core significance. Network analysis was then used to map the relationships among these words, and modularity analysis helped identify natural groupings or clusters. These analytical methods collectively provided a holistic view of research trends, significant concepts, and the overall structure of English writing research in the two periods under scrutiny.

Frequency analysis identified keywords that consistently appeared with high frequency in both periods, such as ‘test’, ‘college students’, ‘difference’, ‘vocabulary’, and ‘questionnaire’, indicating core areas of interest in writing research. However, the 2010s witnessed the emergence of new research fields, including ‘task-based approaches’, ‘textbook analysis’, ‘corpus analysis’, and ‘genre-based approaches’, diversifying and specializing writing research. Centrality analysis, performed on the most frequent words for each period, revealed influential nodes and demonstrated that word frequency does not directly correspond to its centrality within the keyword network. In the 2000s, ‘questionnaire’ played a pivotal role in various studies, while the central position shifted to ‘feedback’ in the 2010s, emphasizing its role in writing instruction. Additionally, the analysis highlighted the growing significance of ‘peer feedback’ as a research interest, underscoring its benefits in enhancing writing skills.

The centrality analyses facilitated the creation of period-specific network maps, illustrating the intricate interconnections among research themes. For instance, in the 2000s, the influential node ‘questionnaire’ was diversely connected, while the 2010s saw a shift towards the ‘feedback’ node, which had consistent ties with ‘error’, exemplifying the persistent focus on error correction. Modularity analysis further divided the network into distinctive communities per period, resulting in easily interpretable maps that displayed groupings such as the ‘sentence group’, ‘classroom group’, and ‘test group’ in period 1, and the ‘relationship group’, ‘error group’, and ‘peer feedback group’ in period 2. These network maps clarified the organization of node clusters and the dynamic evolution of English writing research over the two periods, contributing to a clearer understanding of the interconnections and evolving structure of the field.

The findings of this study significantly contribute to the field of English writing research in Korea. Through a comprehensive analysis of research papers and the utilization of semantic network analysis, key shifts in the focus of English writing education and research between the 2000s and 2010s were identified. The creation of network maps for each period, supported by modularity analysis, provided a visual representation of the interconnections and groupings of research themes, enhancing our understanding of the overall framework of English writing research. These findings shed light on the evolving landscape of the field and offer valuable insights for researchers and practitioners in English writing education in Korea.

However, it is important to acknowledge the limitations of this study. Firstly, it should be noted that the analysis was conducted based on a specific set of research papers collected using Biblio Data Collector. While efforts were made to gather data related to English writing and published from 2001 to 2020 in Korea, it is important to acknowledge that there is a possibility that some papers may have been missed. Therefore, the findings of this study should be interpreted within the context of the selected dataset and may not fully encompass the entirety of English writing research conducted in Korea. Additionally, the study focused solely on the Korean context, and cross-cultural perspectives in English writing research were not explored. Future research could examine similarities and differences in writing research across different cultural and educational contexts. Comparative studies across different countries

and educational systems could shed light on the cultural, institutional, and instructional factors that shape writing research and pedagogy. Such cross-cultural investigations would enable us to identify best practices, share innovative approaches, and promote international collaborations in the field of second language writing.

References

- Chae, S. 2012. Identifying effective English L2 writing interventions: Emerging trends and issues in recent research. *English Teaching* 67(4), 3-24.
- Cumming, A. 1990. Metalinguistic and ideational thinking in second language composing. *Written Communication* 7(4), 395-411.
- Fatimah, F. and H. Masduqi. 2017. Research trends in EFL writing in Indonesia: Where art thou? *Journal of Teaching and Education* 7(1), 89-98.
- Fortunato, S. and M. Barthélemy. 2007. Resolution limit in community detection. *Proceedings of the National Academy of Sciences* 104(1), 36-41.
- Freeman, L. 1977. A set of measures of centrality based on betweenness. *Sociometry* 40(1), 35-41.
- Freeman, L. 1978/79. Centrality in social networks conceptual clarification. *Social Networks* 1(3), 215-239.
- Girvan, M. and M. E. J. Newman. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99(12), 7821-7826.
- Hwang, S. and J. Kim. 2019. Research trends analysis of North Korean English education using keyword network analysis. *Journal of Learner-Centered Curriculum and Instruction* 9(18), 95-120.
- Jun, B. and C. Han. 2013. A method to decide the number of additional edges to integrate the communities in social network by using modularity. *Journal of the Korea Society of Computer and Information* 18(7), 101-109.
- Kamada, T. and S. Kawai. 1989. An algorithm for drawing general undirected graphs. *Information Processing Letters* 31(1), 7-15.
- Kang, D. 2006. Synthesis of 40 years's writing studies in the journal of *English Teaching*. *English Teaching* 61(special issue), 97-114.
- Kim, Y. and Y. Kim. 2019. *Social Network Analysis*. Seoul: Parkyoungsa Co.
- Klein, P. D. and P. Boscolo. 2016. Trends in research on writing as a learning activity. *Journal of Writing Research* 7(3), 311-350.
- Kwon, E. 2021. Research trends in AI-based English language teaching and learning. *Korean Journal of English Language and Linguistics* 21, 1313-1337.
- Li, M. and N. Storch. 2017. Second language writing in the age of CMC: Affordances, multimodality, and collaboration. *Journal of Second Language Writing* 36, 1-5.
- Lim, J. 2020. A comprehensive review of reading- and writing-related articles in *Modern English Education* in its 20 years of history. *Modern English Education* 21(4), 95-111.
- Loncar, M., W. Schams and J. Liang. 2023. Multiple technologies, multiple sources: trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015-2019, *Computer Assisted Language Learning* 36(4), 722-784.
- Ma, J. 2015. Changing trends in English second language writing research in *English Teaching*. *English Teaching* 70(5), 329-353.
- Manchón, R. 2011. *Learning-to-Write and Writing-to-Learn in an Additional Language*. Philadelphia: John Benjamins.

- Medaglia, J. and D. Bassett. 2017. Network analyses and nervous system disorders. *Oxford Research Encyclopedia of Neuroscience*. <https://arxiv.org/abs/1701.01101>
- Moreno, J. L. 1934. *Who Shall Survive?: A New Approach to the Problem of Human Interrelations*. Washington: Nervous and Mental Disease Publishing Co.
- Newman, M. E. J. 2006a. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E* 74, 036104.
- Newman, M. E. J. 2006b. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 103(23), 8577-8582.
- Newman, M. E. J. 2018. *Networks* (2nd ed.). Oxford: Oxford University Press.
- Newman, M. E. J. and M. Girvan. 2004. Finding and evaluating community structure in networks, *Physical Review E* 69, 026113.
- Park, E. 2021. Topic analysis in EFL writing in Korea using text mining. *Korean Journal of Applied Linguistics* 37(3), 95-122.
- Park, H. and S. Jang. 2014. A comprehensive review of English writing studies in *English Language and Literature Teaching* over the past 20 years. *English Language and Literature Teaching* 20(1), 403-426.
- Riazi, M., L. Shi and J. Haggerty. 2018. Analysis of the empirical research in the *Journal of Second Language Writing* at its 25th year (1992–2016). *Journal of Second Language Writing* 41, 41-54.
- Shin, Y. and Y. Kim. 2020. Research and trends in English language education from 1990 to 2019: A keyword analysis of published articles in *English Language Teaching*. *English Language Teaching* 32(4), 205-224.
- Swain, M. 1985. Communicative competence: Some role of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden, eds., *Input in Second Language Acquisition*, 235-253. MA: Newbury House.
- Swain, M. 2000. The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf, ed., *Sociocultural Theory and Second Language Learning*, 97-114. Oxford: Oxford University Press.
- Swain, M. and S. Lapkin. 1995. Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics* 16, 371-391.
- Won, Y. and Y. Kim. 2021. Analysis of research trends in Korean English education journals using topic modeling. *International Journal of Contents* 21(4), 50-59.
- Zhang, J. and Y. Luo. 2017. Degree centrality, betweenness centrality, and closeness centrality in social network. *Advances in Intelligent Systems Research* 132, 300-303.

Examples in: English
Applicable Languages: English
Applicable Level: All