

Phoneme–grapheme recognition in early EFL literacy development: A scientometric and bibliometric mapping study

Elmiati^{1✉}, Nurul Hikmah², Rosdiana³, Linda Ningsih⁴, Nurulakmiyati⁵

¹²³⁴⁵SDN 34 Dompu, Indonesia

¹Email: elmiatiaridyelmiati@gmail.com



(✉ Corresponding Author)

Abstract

Phoneme–grapheme recognition plays a critical role in early literacy development, particularly in English as a Foreign Language (EFL) learning contexts where sound–symbol relationships differ from learners’ first language systems. This study aims to map the development of research on phoneme–grapheme recognition in early EFL literacy through a bibliometric and scientometric analysis of publications indexed between 2019 and 2025. Using data retrieved through Publish or Perish from the Scopus database, a total of 500 articles were identified, with 499 meeting the inclusion criteria for analysis. Descriptive and network-based analyses were conducted using Microsoft Excel, VOSviewer, and Bibliometrix, enabling examination of publication trends, citation patterns, authorship networks, and thematic evolution. Results show a steady increase in publications over the observed period, with the highest growth recorded in 2025. Co-occurrence visualizations revealed four dominant thematic clusters: foundational phonological skills, instructional practices, intervention-based research, and learner-specific challenges, including dyslexia and multilingual contexts. Recent research trends indicate a shift toward digital phonics tools, teacher training, and evidence-based instructional models. This study provides a comprehensive overview of research development in the field and identifies key gaps, including limited longitudinal studies, insufficient focus on neurodiverse learners, and minimal cross-linguistic comparative research. The findings offer a roadmap for future inquiry and evidence-based instructional improvement in early EFL literacy development.

Keywords: phoneme–grapheme recognition, phonics instruction, early literacy, EFL learners, reading development

Citation | Elmiati E., Hikmah N., Rosdiana R., Ningsih L., Nurulakmiyati N. (2025). Phoneme–grapheme recognition in early EFL literacy development: A scientometric and bibliometric mapping study. *Journal of Language and Literacy Learning*, 1(2), 13–28.

Received: December 5, 2025 | **Revised:** December 26, 2025 | **Accepted:** December 27, 2025 | **Published:** December 30, 2025

Contribution to Literature: This study provides the first bibliometric mapping of research on phoneme–grapheme recognition in early EFL literacy, particularly in multilingual contexts, technology-enhanced phonics, and early identification of reading difficulties, thereby establishing a foundation for future empirical studies and scholarly development.

1. Introduction

Early literacy serves as the foundation for children’s later academic success, particularly in contexts where learners engage with an additional language. In English as a Foreign Language (EFL) settings, developing the ability to decode written text is a key milestone in literacy development, as it enables learners to transition from emergent reading behaviors to more advanced processes such as word recognition, fluency, and comprehension. One of the core cognitive components underpinning this transition is phoneme–grapheme recognition (PGR), the ability to connect speech sounds (phonemes) with written symbols (graphemes). Research in first language literacy has long demonstrated that strong PGR and phonological awareness predict reading achievement and support orthographic learning (Castles et al., 2018; Hulme et al., 2020). Likewise, evidence from EFL contexts suggests that explicit instruction in sound–letter correspondence significantly contributes to improved decoding and reading fluency among young learners (Koda & Yamashita, 2018; McBride et al., 2022).

Despite the importance of PGR for reading development, research in this area within EFL environments has expanded unevenly across regions and academic disciplines. While some countries have integrated systematic phonics and phonology-based instruction into primary school language curricula, others particularly in Asia, continue to rely on memorization, whole-word recognition, or textbook-driven instruction with limited explicit emphasis on phoneme–grapheme correspondence (Butler, 2015; Russak & Kahn-Horwitz, 2015). The rapid expansion of English in early education over the past two decades has further intensified interest in foundational literacy research, especially as more schools adopt early English instruction and digital learning tools. However, the extent to which global scholarship has addressed PGR specifically, and how this body of knowledge has evolved over time, remains unclear.

Given this growing interest and the pedagogical relevance of phoneme–grapheme recognition for early literacy development, there is a need to systematically map how the field has developed, which research communities have contributed most, and what conceptual or methodological patterns dominate the literature. Bibliometric and scientometric approaches provide a means of addressing these questions by analyzing publication patterns, thematic clusters, citation structures, and collaboration networks across time. Such approaches not only reveal how a field has evolved but also help identify emerging research fronts, influential scholars, dominant theories, and under-examined areas that require further inquiry.

Therefore, this study aims to provide a comprehensive overview of global research related to phoneme–grapheme recognition in early EFL literacy development. Using bibliometric and scientometric mapping techniques, this research analyzes peer-reviewed publications to explore how scholarship in this domain has grown, which areas have experienced the most development, and where research gaps remain. The study is guided by the following research questions: (1) How has research on phoneme–grapheme recognition in early EFL literacy development evolved across time? (2) Who are the key contributors, publication venues, and research communities shaping this field? (3) What themes, theories, and methodological patterns characterize existing research on phoneme–grapheme recognition in early EFL literacy development? (4) What knowledge gaps and future research opportunities emerge from the current body of literature?

By mapping the intellectual structure and evolution of research in this domain, the present study contributes a synthesized perspective that may inform future research agendas, guide curriculum and assessment development, and support evidence-based practices in early EFL literacy instruction.

2. Literature Review

Research on early literacy development consistently highlights the central role of phonological processing and decoding skills in learning to read, particularly among beginning learners and young children acquiring a new language. In the context of English as a Foreign Language (EFL), phoneme–grapheme recognition (PGR) serves as a foundational component of early literacy because it enables learners to map spoken sounds onto written forms, supporting decoding, spelling, and the development of orthographic knowledge. Studies in first language settings have established strong correlations between explicit PGR instruction and improvements in reading fluency and comprehension (Castles et al., 2018; Ehri, 2020). This body of evidence has influenced instructional models such as systematic phonics, phonological awareness training, and multisensory literacy approaches.

In EFL settings, however, the development of phoneme–grapheme correspondence presents unique challenges due to differences between learners’ native phonological systems and English orthography. Research indicates that learners whose first language has a transparent orthography or different phonemic inventory often experience difficulties recognizing English vowel variations, consonant clusters, and irregular spelling patterns (Koda & Yamashita, 2018). These complexities have encouraged researchers and educators to explore instructional interventions that explicitly teach sound–symbol relationships to support early decoding ability. Recent studies also highlight the increasing role of digital tools, gamified phonics applications, and technology-assisted pronunciation feedback systems in facilitating PGR development (McBride et al., 2022).

Although research related to PGR in early EFL literacy has expanded across pedagogical, linguistic, and technological domains, the literature remains dispersed across multiple fields and lacks a synthesized map of conceptual development, influential scholarship, and dominant methodological trends. Some studies focus on intervention effectiveness, others examine developmental differences, while emerging research explores assessment tools or instructional frameworks. Despite this diversity, there has been no comprehensive bibliometric or scientometric synthesis that examines how knowledge in this area has evolved, who the leading contributors are, and where theoretical or empirical gaps persist.

To position the present study within existing scholarship, Table 1 summarizes selected previous studies relevant to phoneme–grapheme recognition and early EFL literacy. The table highlights core themes, methodologies, and key findings, providing an overview of how research has been approached across different contexts.

Table 1. Research on phoneme–grapheme recognition in early EFL literacy development.

| No | Title | Reference |
|----|--|---------------------|
| 1 | Developing Phoneme–Grapheme Recognition for English as a Foreign Language: A Longitudinal Study at Japanese Primary School | Nakao et al. (2024) |
| 2 | Children’s Language Skills Can Be Improved: Lessons From Psychological Science for Educational Policy | Hulme et al. (2020) |
| 3 | The Science of Learning to Read Words: A Case for Systematic | Ehri (2020) |

| No | Title | Reference |
|----|--|--------------------------|
| 4 | Phonics Instruction Reading and Writing Words: A Cross-Linguistic Perspective | McBride et al. (2022) |

3. Research Method

3.1. Research Design

This study employed a bibliometric and scientometric research design to analyze the development of scholarly work related to phoneme–grapheme recognition in early EFL literacy development. Bibliometric analysis was selected because it enables systematic examination of publication patterns, citation structures, and thematic evolution within a research domain. Scientometric visualization further supports the identification of knowledge clusters, collaboration networks, and trending research themes through network mapping.

The analysis focused on publications indexed from 2019 to 2025. The *Publish or Perish* software was used to retrieve metadata from Scopus, given its strong coverage of peer-reviewed literature in language education, applied linguistics, and reading development. The initial search yielded 500 publications based on predefined search terms. A screening process was conducted to ensure relevance and data completeness. After removing unrelated items, duplicates, and studies lacking analyzable metadata, 499 publications were retained for further analysis. The final dataset was exported in CSV format to support statistical analysis and visualization.

3.2. Data Sources

All bibliographic data used in this study were retrieved from the Scopus database due to its consistency in indexing metadata such as author affiliations, citation counts, keywords, abstracts, and references. The search strategy applied the following Boolean string: ("phoneme-grapheme" OR "sound-letter correspondence" OR "phonics" OR "letter-sound mapping") AND ("early literacy" OR "early reading" OR "emergent literacy") AND ("EFL" OR "English as a foreign language" OR "second language learners")

After retrieval, titles and abstracts were screened to confirm alignment with the study focus. Articles centered solely on general literacy, speech therapy, neurological research, or native-language phonics without an EFL context were excluded. The final dataset represents research specifically addressing phoneme–grapheme recognition or closely related constructs within EFL early literacy.

3.3. Procedure of Data Collection

The data collection process consisted of the following steps: (1) Retrieval of Publications: Metadata were extracted using Publish or Perish from Scopus using the finalized search string. (2) Screening and Filtering: Titles, abstracts, and keywords were reviewed to verify relevance. Duplicate and non-eligible studies were removed. (3) Preparation and Formatting: The dataset was exported in CSV for statistical tabulation in Microsoft Excel and in BibTeX format for analysis in VOSviewer and Bibliometrix (R package). (4) Computational Mapping: VOSviewer was used to generate co-occurrence maps, citation networks, and collaboration visualizations to identify research clusters and thematic structures. (5) Verification: The visual and statistical outputs were reviewed to ensure accuracy and alignment with the study objectives.

3.4. Data Analysis

Data analysis was conducted using a two-step approach. First, descriptive statistics were generated using Microsoft Excel to identify annual publication trends, most productive authors, countries, and journals. This provided an overview of scholarly activity in the field.

Second, scientometric network analysis was performed using VOSviewer and Bibliometrix. These tools enabled examination of: keyword co-occurrence patterns; co-citation and bibliographic coupling; author and institutional collaboration networks; thematic and temporal evolution of research focus.

Link strength, node density, and cluster formation were analyzed to interpret relationships between concepts and identify emerging or underexplored areas. The combination of statistical and visualization approaches allowed this review to not only map the structure of the field but also highlight potential knowledge gaps and future research directions.

4. Results and Discussion

4.1. Phoneme–Grapheme Recognition In Early EFL Literacy Development 2019- 2025

The analysis of publication output from 2019 to 2025 indicates a clear upward trajectory in research focusing on phoneme–grapheme recognition within early EFL literacy development. From the initial dataset of 500 retrieved articles, 499 publications met the inclusion criteria and were retained for bibliometric analysis. The removal of one article was due to incomplete metadata, ensuring the final dataset remained clean, consistent, and analyzable.

As shown in Table 2 and visualized in Figure 1, the earliest publication activity in 2019 recorded 52 documents, marking the beginning of visible interest in this topic. Research activity increased gradually in 2020 with 65 publications, followed by continued growth in 2021, which recorded 76 publications and marked the first notable peak.

Although a slight decline occurred in 2023 with 66 documents, the overall trajectory remained positive. Publication numbers stabilized in 2022 and 2024, each with 74 documents, suggesting sustained scholarly engagement. The most significant surge occurred in 2025, where publication output reached 92 documents, representing the highest point in the seven-year period.

Collectively, these findings demonstrate that research on phoneme–grapheme recognition in early EFL literacy development has expanded consistently, gaining stronger momentum in recent years. This progression reflects increasing global attention to early literacy interventions, phonological instruction, and foundational decoding skills in foreign language learning environments.

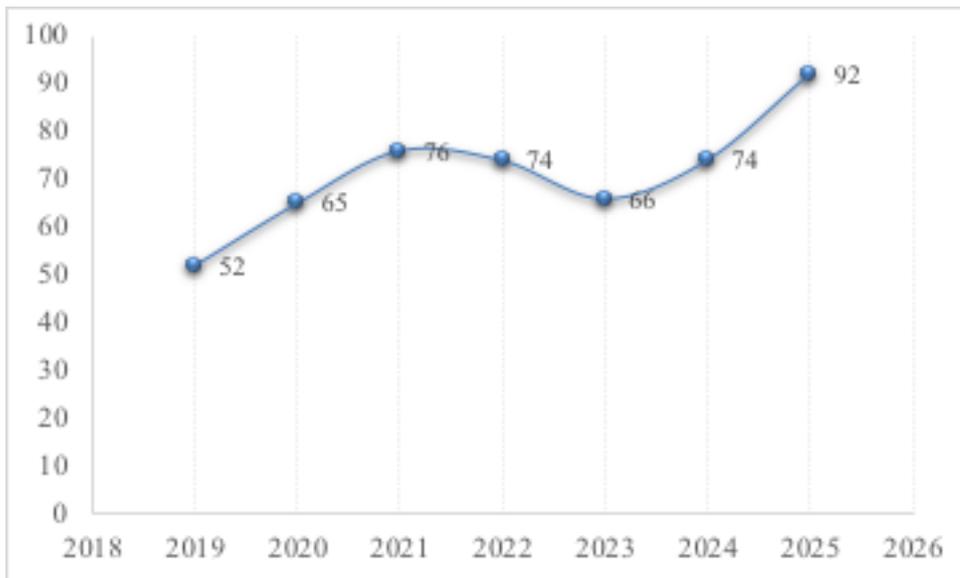


Figure 1. Annual report of publications.

Table 2. Annual report research on “phoneme–grapheme recognition in early EFL literacy development”.

| Years | Documents | Percentages (%) |
|-------|-----------|-----------------|
| 2019 | 52 | 10,42 % |
| 2020 | 65 | 13,02 % |
| 2021 | 76 | 15,23 % |
| 2022 | 74 | 14,82 % |
| 2023 | 66 | 13,22 % |
| 2024 | 74 | 14,82 % |
| 2025 | 92 | 18,43 % |
| Total | 499 | 100 % |

4.2. Most Influential and Highly Cited Publications

The analysis of citation patterns provides insight into the scholarly impact and intellectual foundations shaping research on phoneme–grapheme recognition in early EFL literacy development. As shown in Table 3, citation counts vary significantly, reflecting differences in research focus, methodological rigor, accessibility, and relevance across contexts.

The most highly cited publication in the dataset is *A Study on Vocabulary-Learning Problems Encountered by BA English Majors at the University Level of Education* (Afzal, 2019), with 452 citations. Although not exclusively focused on phoneme–grapheme correspondence, this study is widely referenced due to its emphasis on foundational language acquisition challenges, which aligns with early literacy and decoding skill development in EFL settings.

Additional influential works include Spaul et al. (2020) with 234 citations, addressing early-grade reading benchmarks in multilingual contexts, and Teale et al. (2020) with 112 citations, offering a decade-long synthesis of early literacy research trends. These publications highlight the significance of comparative literacy development frameworks, early reading measurement, and large-scale empirical analysis, suggesting that researchers in the field frequently draw on broader literacy evidence when examining phoneme–grapheme acquisition.

Several publications also reflect growing interest in instructional innovation and technology integration. For example, Cilliers et al. (2022) and Chen et al. (2022), each

with 107 citations, explore teacher professional development and early bilingual learning outcomes, respectively. Similarly, [Liu et al. \(2019\)](#) and [Jiang et al. \(2023\)](#) demonstrate increasing attention to automatic speech recognition and pronunciation-enhancing technologies, signaling a shift toward digital literacy tools designed to support sound–symbol mapping and decoding.

Meta-analyses by [Ludwig et al. \(2019\)](#), [Graham et al. \(2021\)](#), and [Lovett et al. \(2021\)](#) also appear prominently in the dataset, indicating sustained interest in evidence-based interventions targeting both reading and writing difficulties. Notably, several influential studies address multilingual or bilingual learner populations, such as [Kremin et al. \(2019\)](#), reflecting the relevance of cross-linguistic phonological transfer when acquiring English orthography.

Overall, the citation landscape reveals three dominant research orientations: (1) **Foundational literacy and phonological development**, including vocabulary, decoding, and phonemic awareness. (2) **Instructional and technological intervention models**, such as digital phonics, speech recognition, augmented reality, and coaching frameworks. (3) **Large-scale reviews and meta-analyses**, synthesizing empirical findings and informing policy and curriculum design.

These patterns indicate that while phoneme–grapheme research is embedded within broader EFL literacy scholarship, emerging studies increasingly focus on targeted instructional and technological applications that directly support early decoding. This evolution reflects a gradual shift from descriptive challenges toward solution-focused and innovation-driven pedagogical models.

Table 3. Most cited articles on phoneme–grapheme recognition in early EFL literacy development

| No | Cites | Title | Year | Cites Per Year | Ref. |
|----|-------|---|------|----------------|--|
| 1 | 452 | A study on vocabulary-learning problems encountered by BA English majors at the university level of education | 2019 | 75.33.00 | Afzal (2019) |
| 2 | 234 | Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages | 2020 | 46.80 | Spaull et al. (2020) |
| 3 | 112 | Early literacy research, 2006–2015: A decade of measured progress | 2020 | 22.40 | Teale et al. (2020) |
| 4 | 107 | Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning | 2022 | 35.67 | Cilliers et al. (2022) |
| 5 | 107 | A burden or a boost: The impact of early childhood English learning experience on lower elementary English and Chinese achievement | 2022 | 35.67 | Chen et al. (2022) |
| 6 | 98 | Are reading interventions for English language learners effective? A meta-analysis | 2019 | 16.33 | Ludwig et al. (2019) |
| 7 | 97 | Do children with reading difficulties experience writing | 2021 | 24.25.00 | Graham et al. (2021) |

| No | Cites | Title | Year | Cites Per Year | Ref. |
|----|-------|---|------|----------------|--|
| | | difficulties? A meta-analysis. | | | |
| 8 | 91 | Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa | 2019 | 15.17 | Kotze et al. (2019) |
| 9 | 86 | Effective intervention for adolescents with reading disabilities: Combining reading and motivational remediation to improve outcomes. | 2021 | 21.50 | Lovett et al. (2021) |
| 10 | 82 | The effects of Spanish heritage language literacy on English reading for Spanish–English bilingual children in the US | 2019 | 0,58819444 | Kremin et al. (2019) |
| 11 | 74 | What is phonological awareness in L2? | 2019 | 12.33 | Saiegh-Haddad (2019) |
| 12 | 66 | Improving English pronunciation via automatic speech recognition technology | 2019 | 11.00 | Liu et al. (2019) |
| 13 | 64 | TED Talks as a listening resource in the EFL college classroom | 2021 | 16.00 | Al-Jarf (2021) |
| 14 | 64 | Online vocabulary tasks for engaging and motivating EFL college students in distance learning during the pandemic and post-pandemic | 2022 | 21.33 | Al-Jarf (2022) |
| 15 | 58 | The effects of ABRACADABRA on reading outcomes: An updated meta-analysis and landscape review of applied field research | 2020 | 0,5 | Abrami et al. (2020) |
| 16 | 58 | Family literacy coaching: Partnering with parents for reading success. | 2019 | 0,42152778 | Brown et al. (2019) |
| 17 | 57 | Exploring the effectiveness of augmented reality technology on reading comprehension skills among early childhood pupils with learning disabilities | 2024 | 57.00.00 | Shaaban & Mohamed (2024) |
| 18 | 52 | Teaching phonetics to enhance pronunciation in an ESL classroom | 2020 | 10.40 | Priya & Kumar (2020) |
| 19 | 51 | Strategies to improve English vocabulary and spelling in the classroom for ELL, ESL, EO and LD students | 2019 | 08.50 | Alshahrani (2019) |
| 20 | 50 | Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom | 2023 | 25.00.00 | Jiang et al. (2023) |

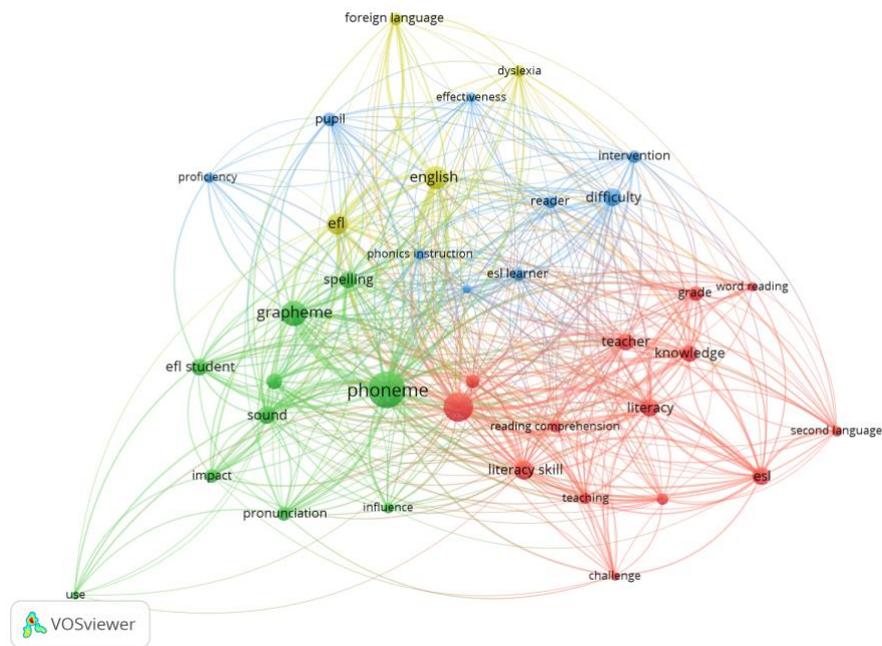


Figure 2. Network visualization based on co-occurrence of terms.

4.3. Visualization of Research Data Mapping of Phoneme–Grapheme Recognition In Early EFL Literacy Development

A bibliometric network visualization was generated using VOSviewer to analyze conceptual relationships between frequently occurring terms within the selected publication set. The process began with 2,795 extracted terms, which were automatically detected across titles, abstracts, and keywords. To ensure analytical relevance, a minimum occurrence threshold of 10 repetitions was applied, reducing the set to 61 terms. From these, the most meaningful terms, representing **60% of the frequency distribution were refined further to a final list of 37 terms, forming the core dataset for network visualization and thematic interpretation.

As illustrated in Figure 2, the co-occurrence network reveals interconnected thematic clusters, each representing a dominant conceptual focus in research related to phoneme–grapheme recognition and early literacy in EFL contexts. Four major clusters were identified.

Cluster 1 (Red) – 14 items: This cluster reflects foundational constructs related to early literacy development. Terms such as *phonemic awareness*, *literacy skills*, *letter recognition*, *reading comprehension*, *word reading*, and *word recognition* suggest a strong emphasis on decoding skills and the role of phonological awareness in supporting early reading acquisition. The presence of terms like *teacher*, *teaching*, and *second language* highlights pedagogical implementation as a central concern.

Cluster 2 (Green) – 10 items: This group focuses on the relationship between *phoneme–grapheme mapping* and its learning outcomes, represented by terms such as *grapheme*, *phoneme*, *pronunciation*, *spelling*, and *comprehension*. The appearance of terms such as *impact*, *influence*, and *use* indicates that researchers are examining instructional effects and learner responses rather than only describing literacy challenges.

Cluster 3 (Blue) – 9 items: This cluster represents intervention-based and applied research orientations. Terms including *intervention*, *phonics instruction*, *letter–sound*, *effectiveness*, *proficiency*, and *pupil* demonstrate interest in instructional frameworks

designed to improve reading performance. The presence of *difficulty* suggests continued attention to barriers faced by early-stage learners, particularly ESL/EFL populations.

Cluster 4 (Yellow) – 4 items: This smaller but significant cluster includes terms such as *dyslexia*, *foreign language*, and *EFL*, indicating a specific research thread exploring phoneme–grapheme challenges among learners with reading difficulties or diverse linguistic backgrounds.

Together, these clusters demonstrate that the field is shaped by three interconnected research directions: (1) foundational literacy skills and cognitive processes, (2) instructional and pedagogical intervention studies, and (3) learner-specific or contextualized literacy challenges, including disabilities and multilingual environments.

The overlay visualization in Figure 3 further shows an evolution in thematic focus across time. Earlier research tended to emphasize core concepts such as *phonemic awareness*, *letter recognition*, and *word reading*, while more recent studies have shifted toward *technology-supported instruction*, *intervention effectiveness*, and *learner variability*. This transition suggests a growing movement from descriptive and diagnostic perspectives toward evidence-based instructional innovation and applied educational technology in EFL literacy development.

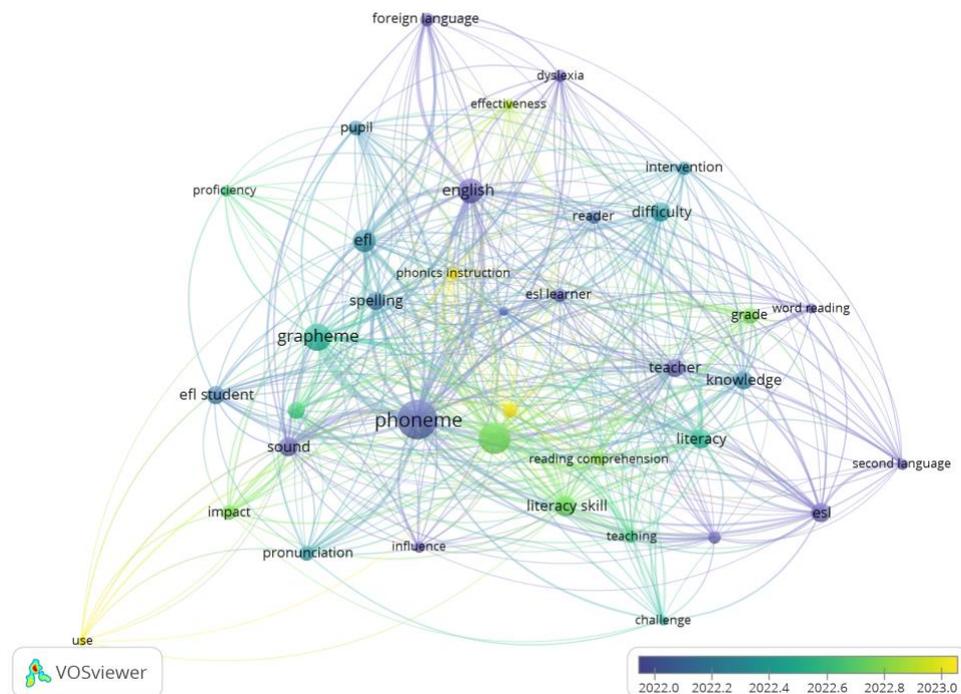


Figure 3. Overlay visualization based on co-occurrence of terms.

In addition to the cluster structure, the density visualization (Figure 4) provides insight into the concentration and prominence of key concepts across the literature. Brighter areas indicate higher frequency and centrality. The term “phoneme” appears as the most dominant focal concept, followed by closely connected terms such as “grapheme,” “spelling,” “literacy,” “phonics instruction,” and “English.” This pattern confirms that research in this domain is anchored primarily in early decoding processes and phonological development. Surrounding terms such as “reading comprehension,” “intervention,” “difficulty,” “pronunciation,” and “teacher” suggest that studies frequently intersect with instructional outcomes and learner challenges.

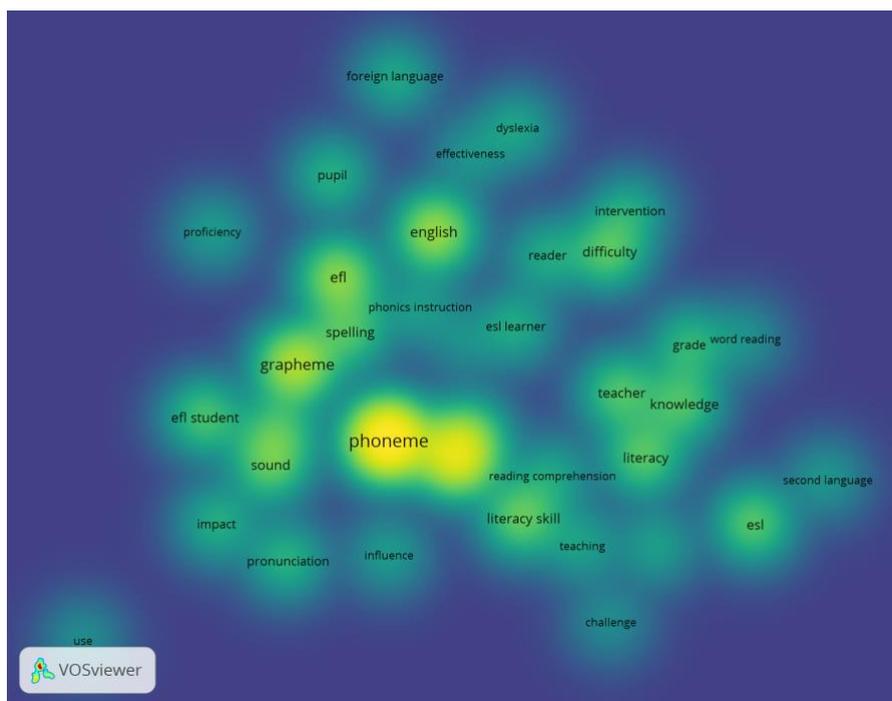


Figure 4. Density visualization based on the co-occurrence of terms.

A more focused examination of individual term networks further clarifies conceptual patterns. The network view centered on the term “phoneme” (Figure 5) demonstrates its strong linkage with both cognitive constructs (e.g., *sound*, *spelling*, *word reading*) and instructional mechanisms (e.g., *letter recognition*, *pronunciation*, and *literacy skill*). This indicates that research draws from both linguistic theory and applied pedagogy when addressing phoneme–grapheme correspondence.

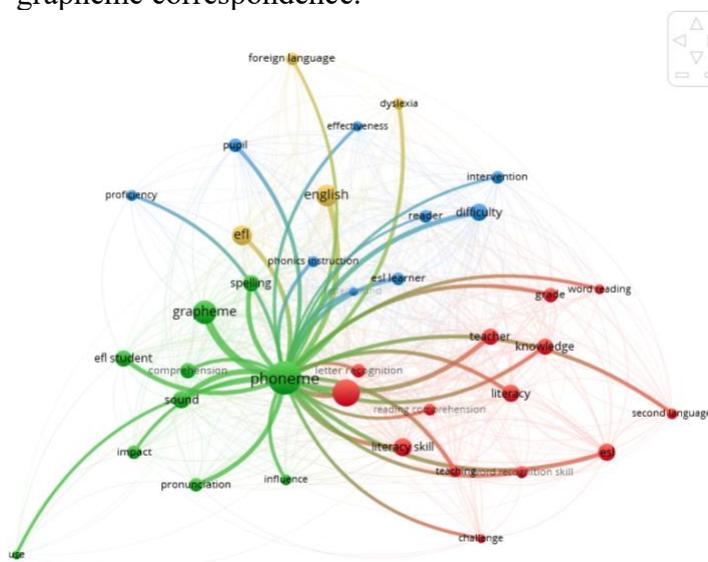


Figure 5. Network visualization of phoneme.

Similarly, the network visualization of “phonics instruction” (Figure 6) highlights its role as an emerging research anchor. Terms such as “intervention,” “effectiveness,” “reader,” “ESL learner,” and “teacher knowledge” appear closely connected, suggesting that recent studies prioritize evaluating instructional models and determining their impact on early literacy achievement in EFL contexts. This reflects a shift from descriptive focus toward empirical validation of classroom-based practices.

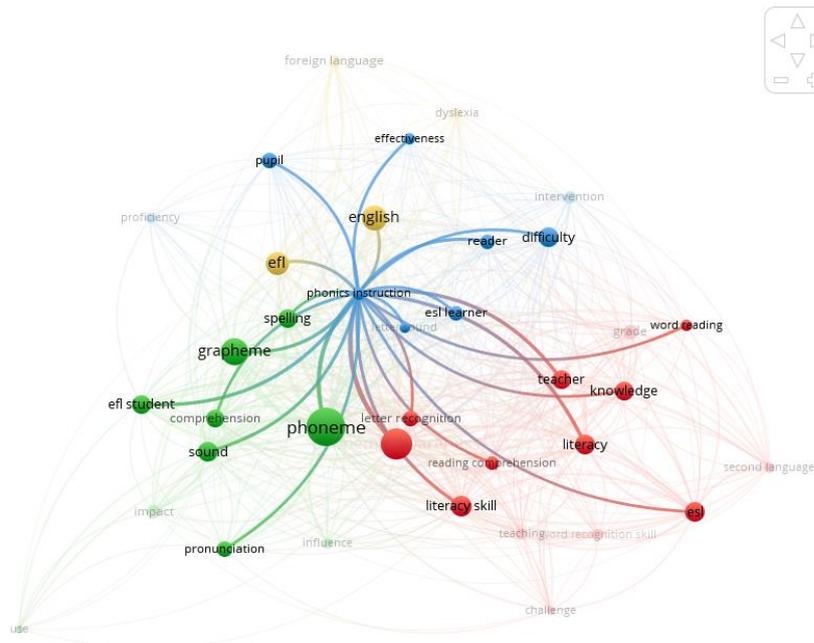


Figure 6. Network visualization of phonics instruction.

Finally, the network mapping centered on the term “EFL” (Figure 7) demonstrates its broad relevance within the field. Connections to “foreign language,” “English,” “grade,” “difficulty,” “dyslexia,” and “teaching” suggest that research consistently considers learner background, linguistic diversity, and instructional conditions as variables that influence phoneme–grapheme development. The presence of terms such as “challenge” and “knowledge” also indicates increasing attention to learner variability and differentiated instruction.

Taken together, these visualizations show a coherent thematic trajectory: research in phoneme–grapheme recognition in early EFL literacy has evolved from foundational cognitive inquiry toward pedagogical applications, intervention evaluation, and context-specific learning challenges. This progression suggests a maturing field increasingly oriented toward practical implementation, teacher expertise, and adaptive literacy support for diverse early learners.

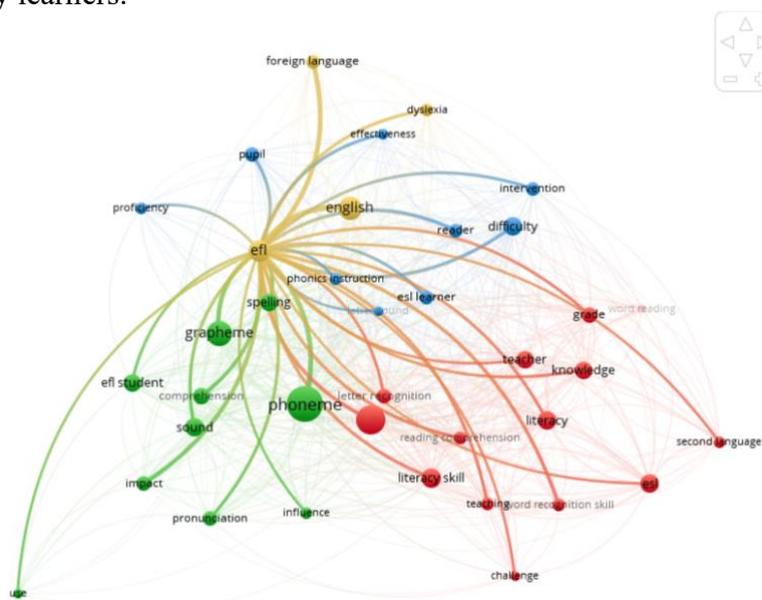


Figure 7. Network visualization of EFL.

4.4. Synthesis and Implications

The upward trajectory of publications between 2019 and 2025 suggests increasing scholarly attention toward phoneme–grapheme correspondence as a core predictor of early literacy success in EFL settings. This finding aligns with previous developmental research showing that phonological awareness and decoding ability serve as the strongest early literacy indicators across language populations (Ehri, 2020). Similarly, the dominance of terms such as *phoneme*, *grapheme*, *phonics instruction*, and *reading comprehension* across the co-occurrence network confirms the discipline’s theoretical grounding in alphabetic principle frameworks, consistent with earlier reviews of phonics-based early literacy instruction (Castles et al., 2018).

The cluster distribution also reflects evolving research priorities. The strong presence of teacher-related keywords (e.g., *teacher knowledge*, *teaching*, *grade*) echoes findings by Huo & Wang (2017), who reported that instructional effectiveness is closely tied to teachers’ phonological and orthographic knowledge, particularly in EFL contexts where phoneme–grapheme rules differ from learners’ L1 orthographies. Additionally, the relatively recent emergence of technology-related terms corresponds with broader trends in literacy education advocating digital scaffolding tools such as AI-based pronunciation feedback, digital flashcards, and speech recognition systems (Chen 2025). However, while the global shift toward technology-supported phonics is visible, the bibliometric mapping reveals that such studies are still fragmented and concentrated in experimental rather than longitudinal designs.

Another critical observation is the relatively small but distinct cluster referencing learning difficulties and dyslexia. This mirrors earlier bibliometric studies in literacy education indicating that research addressing neurodiversity in EFL settings remains limited and underdeveloped (Maunsell, 2020). Given that orthographic irregularity and cross-linguistic interference can amplify decoding challenges, the scarcity of dyslexia-focused EFL phonics studies highlights a meaningful gap for future inquiry.

Taken together, these findings indicate that while foundational research on phoneme–grapheme recognition is well established, the field is transitioning toward applied, contextually responsive, and technology-enhanced models of instruction. This movement points to several emerging priorities: (1) cross-linguistic comparative studies exploring orthographic distance, (2) professional learning models for teachers in low-resource contexts, (3) adaptive and digital phonics tools, and (4) inclusive frameworks addressing diverse learner needs. Thus, the present review contributes not only a structured mapping of the field but also a forward-looking agenda to inform future scholarship and instructional design in early EFL literacy development.

5. Conclusion

This bibliometric analysis provides a comprehensive overview of research trends, thematic patterns, and intellectual structures in the field of phoneme–grapheme recognition within early EFL literacy development from 2019 to 2025. The results demonstrate a clear growth of scholarly attention toward phonological awareness, phonics instruction, and decoding skills as foundational components of early English literacy learning. The dominance of core terms such as *phoneme*, *grapheme*, *phonics instruction*, and *reading comprehension* reflects the field’s continued emphasis on the alphabetic principle and the role of systematic phonics in supporting literacy acquisition.

The clustering and visualization findings highlight four major thematic directions: instructional foundations, learner-related challenges, pedagogical intervention strategies, and broader literacy outcomes. Recent shifts indicate increasing interest in technology-supported approaches and teacher professional competence, signaling a transition from

theoretical and diagnostic work toward applied and implementation-based research. Despite this progress, gaps remain, particularly in research examining neurodiverse learners, longitudinal literacy development, and context-sensitive pedagogical models for multilingual EFL environments.

Overall, this study contributes a structured representation of how phoneme–grapheme research has evolved and where it is heading. By mapping existing evidence and identifying underexplored areas, the findings serve as a foundation for advancing future research and informing instructional development. Continued efforts to integrate phonological science, classroom practice, technological innovation, and inclusive learning frameworks will be essential for strengthening early EFL literacy outcomes globally.

Conflict of Interest Statement

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding Acknowledgment

This research received no external funding from any party.

Author Contributions

Elmiati: Conceptualization, initial draft writing; Hikmah: methodology; Rosdiana: data collection; Ningsih and Nurulakmiyati: editing, and revision.

References

- Abrami, P. C., Lysenko, L., & Borokhovski, E. (2020). The effects of ABRACADABRA on reading outcomes: An updated meta-analysis and landscape review of applied field research. *Journal of Computer Assisted Learning*, 36(3), 260–279. <https://doi.org/10.1111/jcal.12417>
- Afzal, N. (2019). A Study on vocabulary-learning problems encountered by ba english majors at the university level of education. *SSRN Electronic Journal*, 10(3), 81–98. <https://doi.org/10.2139/ssrn.3465990>
- Al-Jarf, R. (2021). TED talks as a listening resource in the efl college classroom. *International Journal of Language and Literary Studies*, 2(3), 256–267. <https://doi.org/10.36892/ijlls.v2i3.691>
- Al-Jarf, R. (2022). Online vocabulary tasks for engaging and motivating efl college students in distance learning during the pandemic and post-pandemic. *International Journal of English Language Studies*, 4(1), 14–24. <https://doi.org/10.32996/ijels.2022.4.1.2>
- Alshahrani, H. A. F. (2019). Strategies to improve english vocabulary and spelling in the classroom for ell, esl, eo and ld students. *International Journal of Modern Education Studies*, 3(2), 65–81.
- Brown, C. L., Schell, R., Denton, R., & Knode, E. (2019). Family literacy coaching: Partnering with parents for reading success. *School Community Journal*, 29(1), 63–86.
- Butler, Y. G. (2015). English language education among young learners in East Asia: A review of current research (2004–2014). *Language Teaching*, 48(3), 303–342. <https://doi.org/10.1017/S0261444815000105>
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19(1), 5–51. <https://doi.org/10.1177/1529100618772271>

- Chen, S., Zhao, J., De Ruiter, L., Zhou, J., & Huang, J. (2022). A burden or a boost: The impact of early childhood English learning experience on lower elementary English and Chinese achievement. *International Journal of Bilingual Education and Bilingualism*, 25(4), 1212–1229. <https://doi.org/10.1080/13670050.2020.1749230>
- Chen, X. (2025). Enhancing interactive english learning through embedded artificial intelligence technology. *International Journal of High Speed Electronics and Systems*, 2540727. <https://doi.org/10.1142/S0129156425407272>
- Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, N., Taylor, S., & Thulare, T. (2022). Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning. *Journal of Development Economics*, 155, 102815. <https://doi.org/10.1016/j.jdeveco.2021.102815>
- Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly*, 55(S1), S45–S60. <https://doi.org/10.1002/rrq.334>
- Graham, S., Aitken, A. A., Hebert, M., Camping, A., Santangelo, T., Harris, K. R., Eustice, K., Sweet, J. D., & Ng, C. (2021). Do children with reading difficulties experience writing difficulties? A meta-analysis. *Journal of Educational Psychology*, 113(8), 1481–1506. <https://doi.org/10.1037/edu0000643>
- Hulme, C., Snowling, M. J., West, G., Lervåg, A., & Melby-Lervåg, M. (2020). Children’s language skills can be improved: lessons from psychological science for educational policy. *Current Directions in Psychological Science*, 29(4), 372–377. <https://doi.org/10.1177/0963721420923684>
- Huo, S., & Wang, S. (2017). The effectiveness of phonological-based instruction in english as a foreign language students at primary school level: A Research Synthesis. *Frontiers in Education*, 2, 15. <https://doi.org/10.3389/educ.2017.00015>
- Jiang, M. Y., Jong, M. S., Lau, W. W., Chai, C., & Wu, N. (2023). Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom. *Journal of Computer Assisted Learning*, 39(1), 125–140. <https://doi.org/10.1111/jcal.12732>
- Koda, K., & Yamashita, J. (2018). *Reading to learn in a foreign language: An integrated approach to foreign language instruction and assessment*. Routledge.
- Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. *International Journal of Educational Development*, 66, 203–213. <https://doi.org/10.1016/j.ijedudev.2018.09.004>
- Kremin, L. V., Arredondo, M. M., Hsu, L. S.-J., Satterfield, T., & Kovelman, I. (2019). The effects of Spanish heritage language literacy on English reading for Spanish–English bilingual children in the US. *International Journal of Bilingual Education and Bilingualism*, 22(2), 192–206. <https://doi.org/10.1080/13670050.2016.1239692>
- Liu, X., Xu, M., Li, M., Han, M., Chen, Z., Mo, Y., Chen, X., & Liu, M. (2019). Improving English pronunciation via automatic speech recognition technology. *International Journal of Innovation and Learning*, 25(2), 126. <https://doi.org/10.1504/IJIL.2019.097674>
- Lovett, M. W., Frijters, J. C., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2021). Effective intervention for adolescents with reading disabilities: Combining reading and motivational remediation to improve outcomes. *Journal of Educational Psychology*, 113(4), 656–689. <https://doi.org/10.1037/edu0000639>

- Ludwig, C., Guo, K., & Georgiou, G. K. (2019). Are reading interventions for English language learners effective? A meta-analysis. *Journal of Learning Disabilities*, 52(3), 220–231. <https://doi.org/10.1177/0022219419825855>
- Maunsell, M. (2020). Dyslexia in a global context: A cross-linguistic, cross-cultural perspective. *Latin American Journal of Content & Language Integrated Learning*, 13(1), 92–113. <https://doi.org/10.5294/laclil.2020.13.1.6>
- McBride, C., Pan, D. J., & Mohseni, F. (2022). Reading and writing words: A cross-linguistic perspective. *Scientific Studies of Reading*, 26(2), 125–138. <https://doi.org/10.1080/10888438.2021.1920595>
- Nakao, K., Oga-Baldwin, Q. W. L., & Fryer, L. K. (2024). Developing phoneme-grapheme recognition for English as a foreign language: A longitudinal study at Japanese primary school. *International Electronic Journal of Elementary Education*, 16(3), 337–347. <https://doi.org/10.26822/iejee.2024.335>
- Priya, M. L. S., & Kumar, P. N. S. (2020). Teaching phonetics to enhance pronunciation in an ESL classroom. *Journal of Critical Reviews*, 7(02), 669–672. <https://doi.org/10.31838/jcr.07.02.121>
- Russak, S., & Kahn-Horwitz, J. (2015). English as a foreign language spelling: Comparisons between good and poor spellers. *Journal of Research in Reading*, 38(3), 307–330. <https://doi.org/10.1111/jrir.12009>
- Saiegh-Haddad, E. (2019). What is phonological awareness in L2? *Journal of Neurolinguistics*, 50, 17–27. <https://doi.org/10.1016/j.jneuroling.2017.11.001>
- Shaaban, T. S., & Mohamed, A. M. (2024). Exploring the effectiveness of augmented reality technology on reading comprehension skills among early childhood pupils with learning disabilities. *Journal of Computers in Education*, 11(2), 423–444. <https://doi.org/10.1007/s40692-023-00269-9>
- Spaull, N., Pretorius, E., & Mohohlwane, N. (2020). Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages. *South African Journal of Childhood Education*, 10(1), 1–14. <https://doi.org/10.4102/sajce.v10i1.773>
- Teale, W. H., Whittingham, C. E., & Hoffman, E. B. (2020). Early literacy research, 2006–2015: A decade of measured progress. *Journal of Early Childhood Literacy*, 20(2), 169–222. <https://doi.org/10.1177/1468798418754939>