Department of Education, University of Oxford

A Systematic Review into the Effectiveness of Teaching Phonics in English as a Foreign Language

Eman Albalooshi

Dissertation submitted to the University of Oxford in partial fulfilment of the degree Master of Science in Applied Linguistics & Second Language Acquisition

St Cross College

Trinity Term 2024

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Hamish Chalmers, for believing in the importance of my research and for providing invaluable and insightful feedback throughout this journey. I am also profoundly grateful to Victoria Murphy, who played a pivotal role in helping me embark on my journey at Oxford University.

A heartfelt thanks goes to my incredible team at iLearn Academy, whose dedication and tireless work brought my ideas to life in our classrooms. I am also deeply thankful to my husband for his unwavering support and belief in me, and to my family for their encouragement and love.

Lastly, I extend my appreciation to my ALSLA classmates, each of whom has inspired me with their unique journeys. Your passion and commitment have enriched this experience beyond words.

Abstract

This study investigates the effectiveness of phonics instruction in teaching English as a Foreign Language (EFL). The literature suggests that learning English as a second or additional language presents significant barriers that can hinder learners from achieving proficiency. While there is evidence that phonics instructions can be effective for developing literacy skills, available evidence rarely synthesises findings from interventions that are tailored around EFL learners. A systematic review of 25 studies was conducted, involving participants from diverse linguistic backgrounds. The methodology adhered to the PRISMA guidelines, with studies selected based on strict inclusion criteria that focused on decoding, fluency, and reading comprehension outcomes. The results reveal that phonics instructions can be effective for EFL learners even in non-anglophone contexts. However, certain factors need to be considered such as developing listening comprehension, utilising innovative approaches such as technology tools, and tailoring systematic phonics instructions around the learners' needs. These factors are explored through the lens of the Simple View of Reading model. Findings also suggest that learners whose L1 uses a Roman script generally show larger effect sizes in phonological awareness and reading fluency after phonics instruction, compared to learners from non-Roman script backgrounds. These differences are interpreted through the lens of the Script Distance Hypothesis, which underscores the importance of considering the linguistic distance between L1 and English when designing phonics interventions. The findings suggest that while phonics instruction is broadly effective, it must be tailored to address the specific challenges faced by EFL learners to maximize its impact.

Table of Contents

List of Tables

List of Figures

1. Introduction

Fueled by globalization, the interest in learning and speaking English as a foreign language (EFL) has surged dramatically over the last decade. The literature suggests that learning English as a second or additional language presents significant barriers that can hinder learners from achieving proficiency (Lipka *et al.,* 2005; Lipka & Siegel, 2007). Jamaludin et al. (2016) argue that poor instruction and limited exposure to the language are two strong factors leading learners to fall behind in their acquisition of English and reading. Consequently, EFL students often face substantial challenges in achieving excellence in language acquisition, challenges compounded by limited language exposure and, in many cases, poor quality instruction. Persistent stagnation in standardized test scores has placed literacy achievement at the forefront of national educational debates (Dilgard *et al.,* 2022). Research in applied linguistics has suggested various strategies to improve early literacy skills for EFL learners. However, while these efforts demonstrate a commitment to utilizing research to guide educational practices, it is crucial to recognize that the effectiveness of specific intervention programs can vary depending on the characteristics of the participants (Tharp, 1991). This variability complicates the identification of programs and strategies likely to be effective for specific groups, such as EFL learners. Numerous initiatives, such as What Works Clearinghouse, have attempted to compile intervention research for educators. However, it remains essential to consider the characteristics of study participants when assessing the applicability of results to the student groups one aims to support. Students from culturally and linguistically diverse backgrounds may respond differently to various instructional and intervention programs. Therefore, careful attention must be paid to whether research has substantiated the effectiveness of interventions for specific EFL groups.

2. Background

Research has suggested various components for successful intervention programs for EFL learners. The most comprehensive synthesis of research on reading instruction for students in general comes from the National Reading Panel (2000), which outlined five major components of reading: phonemic awareness, phonics, fluency, comprehension, and vocabulary. However, the NRP report did not distinguish EFL learners as students with unique language needs, nor did it address whether these five components are equally important for EFL learners as for non-EFL learners. Six years later, the National Literacy Panel released a report that specifically examined literacy in EFL learners (August & Shanahan, 2006), suggesting that phonemic

awareness, phonics, fluency, vocabulary, and comprehension are indeed beneficial instructional components for these learners. However, the reviewed interventions in the report tended to produce smaller effect sizes for EFL learners, often requiring modifications to be relevant and beneficial (August *et al.*, 2009). Since this review was limited to publications before 2004, few research articles met the criteria, and it made no clear distinction between phonemic awareness and phonics, despite the former primarily involving auditory skills and the latter being specifically related to literacy skills.

Cross-sectional studies suggest an association between phonological awareness and reading success. According to the National Inquiry into the Teaching of Literacy in Australia (Education Science and Training, 2005) and the Independent Review of the Teaching of Early Reading in the UK (Rose, 2006), phonological awareness is a crucial component of successful reading programs for young learners. Phonics has emerged as the leading strategy for instructing early literacy in English-speaking countries such as the UK, the US, and Australia, with strong support from educational policies. Several lines of evidence affirm that phonics instructions have successfully helped young learners who are learning English as a foreign language and struggling with reading to achieve reading excellence (Dixon *et al.*, 2011; Yeung *et al.*, 2013; Jamaludin *et al.*, 2016). Despite substantial scientific consensus on the subskills necessary for reading, the methods for effectively teaching these subskills are not as well-defined, particularly for teachers managing classrooms with diverse backgrounds and interests while also fostering academic and social skills (Hindman *et al.*, 2020). Teachers remain uncertain about which teaching practices are most effective and how to implement them in specific classroom environments (Washburn *et al.*, 2011). Therefore, a clear review of phonics intervention practices for EFL learners is needed.

Recent research suggests a relationship between early reading skills (phonics and phonemic awareness) and later reading outcomes for EFL learners (Gersten *et al.*, 2007). However, studies have not typically investigated whether increased instruction in phonics and phonemic awareness leads to improvements in reading achievement among EFL learners, nor have they examined whether such instruction enhances broader reading comprehension skills. Since the initial NRP report (2000), literacy education has undergone significant changes, including the introduction of the Common Core State Standards (2010). These standards have facilitated the vertical alignment of reading instruction, particularly enhancing vocabulary and comprehension strategies. The standards emphasize three primary shifts: (1) "regular practice

with complex texts and their academic language," (2) "reading, writing, and speaking grounded in evidence from texts," and (3) "building knowledge through content-rich nonfiction" (CCSSO, 2010). These shifts may have diverted focus away from phonics in teaching practices. However, the recent podcast "Sold a Story" critically examines the widespread reliance on ineffective reading instruction methods in U.S. schools, emphasizing the significant benefits and necessity of incorporating phonics-based approaches. It highlights resistance within the education system to adopting scientifically supported phonics instructions that could dramatically improve literacy outcomes.

2.1.The Simple View of Reading

The Simple View of Reading (SVR) model is a well-established and widely used framework in L1 reading instruction, recently examined in scholarly literature as a model for understanding variations in reading acquisition success and recommending instructional interventions (Sparks, 2021). Sparks (2021) reviews research on L1 and L2 reading that supports the SVR, illustrating how the model can assess foreign language (FL) reading abilities, distinguish between proficient and struggling FL readers, and pinpoint specific strengths and weaknesses. The SVR model posits that reading comprehension is the product of two key components: word decoding and linguistic comprehension. Proficient reading requires both the ability to decode written words and understand spoken language. Phonics instruction is vital in this context because it directly supports word decoding, the foundational skill for reading comprehension. Decoding involves not just recognizing words but understanding how print maps onto speech sounds, which is essential for reading fluency and comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). Empirical evidence supports the SVR by showing that differences in reading comprehension are closely linked to decoding skills, particularly in the early stages of reading development. For instance, Shankweiler et al. (1999) found that decoding ability significantly predicts reading comprehension among young readers. While decoding skills explain more variance in reading comprehension in the early grades, linguistic comprehension becomes increasingly important as students progress (Catts *et al.*, 2006; Spencer & Wagner, 2017). This evidence underscores the importance of phonics instruction in developing decoding skills, which are crucial for effective reading comprehension. Without solid decoding skills, students may struggle with reading comprehension, even if they have strong language comprehension abilities (Sparks, 2021). Thus, the SVR model could serve as a theoretical framework for this review, helping to identify at-risk students and providing insights into structuring instructional interventions.

2.2.Immerging Trends in Delivering Phonics Instructions

The integration of technology in phonics instruction has emerged as a significant innovation, offering new opportunities to enhance early literacy skills. Digital tools, such as interactive apps and educational software, provide engaging, adaptive, and personalized learning experiences that can cater to diverse learner needs. Research indicates that technology-enhanced phonics instruction can be particularly effective in reinforcing phonemic awareness and decoding skills, critical components of early reading development. For instance, McCarthy et al. (2018) found that students who used phonics-based educational apps demonstrated significant improvements in their phonological awareness and reading fluency compared to those who received traditional instruction alone. Similarly, technology can provide immediate feedback, allowing learners to correct errors in real time, which is crucial for mastering phonics rules (Hall *et al.*, 2019). However, while technology offers promising benefits, it is essential to consider its implementation carefully to avoid over-reliance on digital tools at the expense of teacher-led instruction. Studies suggest that the most effective phonics programs integrate technology with traditional methods, ensuring that digital tools complement, rather than replace, the foundational elements of phonics instruction (Higgins & Xiao, 2020). Therefore, while technology can enhance phonics instruction, it should be used strategically within a broader, balanced literacy program to maximize its potential benefits. Thus this review will not exclude studies that implement the use of technology in delivering phonics instructions to report on the available evidence.

Translanguaging refers to the dynamic process by which bilingual or multilingual individuals use their entire linguistic repertoire to make meaning, communicate, and learn, rather than strictly separating their languages. This concept challenges the traditional view of bilingualism as the use of two separate languages, instead emphasizing the fluid and flexible ways in which languages interact in the minds of multilingual speakers. García (2009) describes translanguaging as a practice that goes beyond mere code-switching, allowing individuals to draw on all their linguistic resources in a flexible and integrated manner to enhance understanding and communication. In the context of phonics instruction, translanguaging can be a powerful pedagogical tool, especially for learners of English as a second language (ESL). Utilizing students' home languages alongside English during phonics lessons can help scaffold learning and make connections between the phonological systems of their L1 and L2. Research by García and Wei (2014) highlights how translanguaging practices

can support literacy development by allowing students to access and apply their existing linguistic knowledge, thereby reducing cognitive load and facilitating the acquisition of new phonemic patterns in English. For instance, during a phonics lesson, an educator might encourage students to compare and contrast the sounds of phonemes in their home language with those in English. This approach not only validates the students' linguistic identities but also deepens their understanding of how sounds operate across languages. This method has been shown to enhance phonological awareness, a critical component of successful phonics instruction (Velasco & García, 2014). By integrating translanguaging into phonics instruction, educators can create a more inclusive and effective learning environment that leverages the full linguistic repertoire of bilingual students, ultimately leading to improved literacy outcomes (García *et al.*, 2017). This review will explore the role of phonics when delivered through a translanguaging approach.

2.3.The Script Distance Hypothesis

The script distance hypothesis, when applied to foreign language learning, suggests that the ease with which learners acquire reading skills in a foreign language is influenced by the degree of similarity between the writing systems of their L1 and L2. A smaller script distance, as seen between English and Spanish, where both use the Latin alphabet, allows learners to transfer phonological and orthographic knowledge more easily from their L1 to the L2 (Koda, 2005). In contrast, a larger script distance, such as between English and Chinese, presents significant challenges, as learners must adapt to a completely different orthographic and phonological system, which can impede the effectiveness of phonics instruction and overall reading acquisition (Wang, Koda and Perfetti, 2003). Empirical studies substantiate the impact of script distance on foreign language learning. For example, Wang et al. (2003) found that Chinese learners of English, who are accustomed to a logographic script, experienced greater difficulties in developing phonological awareness in English compared to learners whose L1 is alphabetic, such as Korean or Russian. This difficulty arises because learners must acquire new phoneme-grapheme correspondences that do not exist in their L1. Similarly, Bassetti (2006) observed that Italian learners of English, despite both languages using the Latin script, struggled with English's deeper orthography, which is less transparent than Italian's, thus complicating the application of phonics rules. These studies highlight the significant role that script distance plays in shaping the foreign language learning experience, particularly in reading acquisition.

2.4.Previous Phonics Systematic Reviews

Several studies have affirmed that phonics intervention programs successfully help young learners learning English as a second language, especially those struggling with reading, to achieve reading excellence (Dixon, Schagen and Seedhouse, 2011; Yeung, Siegel and Chan, 2013; Jamaludin *et al.,* 2016). A recent study in Bahrain conducted by Albalooshi (2022) investigated the role of a phonics intervention program in developing reading skills for students facing difficulties. The study found that synthetic phonics instructions significantly helped the experimental group outperform the control group on standardized reading measures. T-test statistics indicated that having different teachers deliver the phonological awareness program did not affect the improvement level. A recent systematic review by Dilgard, Hodges, and Coleman (2022) proposes a strong relationship between phonics instruction interventions and students' attainment of reading fluency. The review identifies several effective strategies within these interventions: (a) explicit instructions benefit struggling students, (b) a balanced approach to teaching both encoding and decoding skills yields favourable results, (c) the specific phonics requirements of different student groups necessitate targeted interventions, (d) intervention programs should align with educational goals, and (e) literacy leaders can effectively utilize different staff members for interventions if they receive adequate support (Dilgard *et al.*, 2022). These strategies are carefully considered in the design of this study's proposed intervention.

However, recent reviews do not necessarily distinguish between interventions based on phonemic awareness and phonics instructions. While both phonemic awareness and phonics are vital components of literacy instruction, they focus on different aspects of reading skills. Phonemic awareness, defined as the ability to hear, identify, and manipulate individual sounds or phonemes in spoken language, is a crucial auditory skill for understanding the sound structure of words, a prerequisite for learning to read (Adams, 1990; Stanovich, 1993). Phonics instruction, on the other hand, involves teaching the relationships between these phonemes and their corresponding letters or letter combinations in written language. It emphasizes decoding skills, enabling children to translate written words into spoken language by sounding out letters and blends (NRP, 2000; Ehri, 2005). While phonemic awareness lays the foundation for understanding how language works at a sound level, phonics provides the tools to apply this understanding to reading and writing. The SVR model also posits that starting reading is not just about recognizing words but understanding how print maps onto speech sounds, essential for reading fluency and comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990).

This review will therefore focus specifically on studies involving phonics instruction to evaluate its role in helping learners utilize it as a tool to begin reading.

2.5.Comparative Analysis of Previous Systematic Reviews

The review by Dilgard et al. (2022) comprehensively examines the impact of phonics instruction on early literacy development, particularly in word recognition and overall reading ability in young learners. It underscores the effectiveness of systematic phonics instruction, especially when combined with phonemic awareness, in enhancing literacy outcomes. However, the geographical focus of the reviewed studies—primarily Anglophone countries—limits the generalizability of the findings to non-Anglophone or multilingual settings. Additionally, the review does not extensively explore the technological aspects of phonics instruction, such as the use of digital tools or online platforms. The studies included in this review span several years, concentrated mostly in the past two decades, reflecting contemporary practices in early literacy education. Although the effect sizes reported are generally positive, indicating the benefits of phonics instruction, the variability in study methodologies and the absence of longitudinal studies weaken the ability to draw definitive conclusions. Moreover, while the review highlights traditional measures of phonics instruction, such as assessments of word decoding and recognition, it offers limited discussion on the specific types of tests or measures used across studies. Importantly, the review does not consistently account for learners' first language (L1) when making comparisons, a significant oversight given the impact that L1 can have on the effectiveness of phonics instruction.

In contrast, the meta-analysis by Murphy Odo (2021) focuses on the effects of phonological awareness (PA) and phonics instruction on word and pseudo-word reading among learners of English as a second language (L2). The studies included in this meta-analysis are more geographically diverse, incorporating research from both Anglophone and non-Anglophone contexts, providing a broader understanding of how phonics instruction functions in different linguistic environments. However, like the Dilgard et al. review, the technological aspects of phonics instruction are not a primary focus, and there is limited exploration of how digital tools or resources might influence learning outcomes. Murphy Odo’s meta-analysis covers studies conducted over nearly three decades, from 1990 to 2019, offering a comprehensive view of the evolution of phonics instruction in L2 settings. The reported effect sizes are moderate to large, with a moderate effect on word reading ($g = 0.53$) and a large effect on pseudo-word reading ($g = 1.51$), which are higher than those typically reported in the

Dilgard et al. (2022) review. This difference may be due to the inclusion of L2 learners and the specific focus on decoding skills, which are more directly impacted by phonics instruction. The meta-analysis also notes significant variability in the measures used across the studies, with some using standardized assessments and others relying on researcher-designed tools. These differences contribute to the variability in reported effect sizes. A critical strength of Murphy and Odo's meta-analysis is its attention to the learners' L1 when analyzing the effectiveness of phonics instruction. The authors consider how the learners' L1 writing system—whether alphabetic, logographic, or another type—might influence their ability to benefit from phonics instruction. This consideration is particularly important in L2 contexts, where the transfer of literacy skills from L1 to L2 can significantly impact learning outcomes. The meta-analysis suggests that learners from different L1 backgrounds may experience varying degrees of success with phonics instruction, highlighting the need for tailored approaches based on learners' linguistic backgrounds. However, this review did not distinguish between phonemic awareness and phonics instructions despite the former primarily involving auditory skills and the latter being specifically related to literacy skills. In summary, while both reviews highlight the effectiveness of phonics instruction, they differ in their scope, geographical focus, and attention to technological and linguistic factors. Dilgard et al. (2022) concentrate primarily on Anglophone contexts and do not extensively explore technological or L1 considerations, whereas Murphy Odo (2021) provides a more geographically and linguistically diverse perspective, with a stronger focus on how learners' L1 influences the effectiveness of phonics instruction. This review addresses these gaps by incorporating a broader range of geographical contexts, paying more attention to the role of technology in phonics instruction, and considering the impact of learners' L1 in both Anglophone and non-Anglophone settings. Additionally, by standardizing the operationalization of key constructs and advocating for longitudinal studies, this review aims to provide a more comprehensive and reliable understanding of the long-term effects of phonics instruction. To our knowledge, this study represents the first systematic effort to explore and evaluate the quality of existing research regarding the influence of phonics teaching on fluency and reading comprehension in English as a foreign language. This review also seeks to assess the effectiveness of phonics instruction, focusing on fluency, decoding and reading comprehension outcomes, and analyzing the effects across diverse learner profiles and settings.

3. Methodology

This systematic review adheres to the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) reporting guidelines.

3.1.Review Question

- What are the effects of teaching phonics to learners of English as a foreign language on decoding and/or fluency and/or reading comprehension?

3.2.Protocol Registration

The protocol for this systematic review was prospectively registered on IDESR (International Database of Education Systematic Reviews) (https://idesr.org/). It can be viewed at https://idesr.org/article/IDESR000142.

3.3.Eligibility Criteria

Table 1 presents the eligibility criteria. Published papers in English and Arabic were included to seek available evidence dealing with language learners in preschool, primary and secondary school contexts worldwide.

3.4.Search Strategy

Databases in two different languages were searched (English, and Arabic). English databases included: The Web of Science Social Sciences Citation Index (SCCI), SCOPUS, British Education Index, PsycInfo, Australia Education Index, and ProQuest Social Science Premium Collection. Arabic databases included: e-Marefa (The Digital Arabic Database), Almanhal (Authoritative Arabic Scholarly Content), Masader (Omani digital research resource tailored to the academic community), and Al Masdar (Arab World Research Source). To develop the search terms, the researcher initially compiled a list of relevant keywords. We devised two search strings to encompass literature in English and Arabic. The Arabic search string underwent review by an Arabic-speaking lecturer at The University of Bahrain. The terms were thoroughly tested through a pilot phase. Additionally, comprehensive testing of both search strings was carried out. For the preliminary search, each search string was utilized in databases corresponding to its language. All search terms were applied within the ABSTRACT search frame, as preliminary testing showed that this approach yielded the most relevant results. The search was conducted in May 2024.

Table 1 Eligibility criteria.

Item	Inclusion Criteria	Rationale
Bibliographic information	**Include 1 –** studies with a complete reference and sufficient information **Exclude 1 –** studies without a complete reference or sufficient information	Verification, traceability, transparency, and ethical considerations. Without sufficient bibliographic information, it would be unfeasible to retrieve the work.
English FL phonics intervention	**Include 2 –** studies that focus on the teaching of phonics (letter-sound correspondence), and teaching of English as a foreign language, or a second language, or and additional language. **Exclude 2 –** studies that may include some aspects of reading intervention, but do not have phonics as a core element; exclude studies of English L1 phonics	To our knowledge, there have been no systematic reviews, meta-analyses or scoping reviews published that have investigated both English and Arabic databases in the area of teaching of English FL phonics.
Outcomes	**Include 3 –** studies with measures of decoding, and/or fluency, and/or reading comprehension in English FL are included **Exclude 3 –** studies with other language knowledge or language proficiency measures (e.g., effect of L1 literacy), affective responses (e.g., attitudes, engagement, motivation), non-linguistic outcomes (e.g., working memory or other social outcomes)	The aim of the review is to inform about the effects of teaching phonics when it comes to reading-related skills and how pedagogical implications can be drawn to improve language proficiency for English FL learners.
Study design	**Include 4 –** empirical studies with randomised controlled trial design, non-equivalent groups designs, matched-pairs designs, and regression discontinuity designs. **Exclude 4 –** case studies, ethnographies, and cross-sectional designs; exclude opinion pieces and analyses of policy documents; exclude observational studies.	The aim of the review is to provide a high level of evidence on the effectiveness of phonics instructions, including randomized controlled trials (RCTs) and quasi-experimental designs might be prioritized due to their stronger potential to establish causality.
Participants	**Include 5 –** studies with participants who are typical children and/or adult learners or learners with learning disabilities (e.g., dyslexia, dyscalculia, hyperlexia) **Exclude 5 –** studies where the participants are exclusively learners with developmental disorders (e.g., Autism Spectrum Disorder).	This follows the suggestions of the SVR model (Gough and Tunmer, 1986) that learners who have poor decoding and/or linguistic comprehension are classified as dyslexic or hyperlexic. It is assumed that phonics instructions can help develop these skills in learners. This review can serve as evidence for/against this.
Publication status	**Include 6 –** published literature or published proceedings **Exclude 6 –** grey literature such as theses and dissertations or unpublished proceedings	This review is limited by time and space, and its update will extend the inclusion criteria to include grey literature.
Language of publication	**Include 7 –** studies that are written in in English or Arabic **Exclude 7 –** studies written in any other language	This review is limited by time and scope. Any update will extend the inclusion criteria to include studies in any other languages and will draw on colleagues proficient in those languages to help screen, extract data, and synthesise.

Table 2 Search strategy.

Language	(1) FL nature of studies	(2) nature of instruction	3) type of intervention	4) study design
English	(“second language*” OR SL OR ESL OR “foreign language*” OR FL OR “modern language*” OR “modern foreign language*” OR MFL OR “additional language*” OR L2 OR EFL)	AND (teach* OR learn* OR instruct* OR pedagog* OR acqui* OR train* OR study* OR educat* OR program*)	AND (phon* OR phonological decoding OR phonological recoding OR phonological coding OR decoding)	AND (experiment* OR quasi-experiment* OR intervention OR RCT OR randomized control* trial OR randomised control* trial OR Regression Discontinuity Design OR RDD OR compar* OR evaluat*)
Arabic	("لغة أجنبية" OR "لغة إضافية" OR "ثانية" OR مكتسبة OR "للناطقين بها" OR "لغير الناطقين بغيرها")	(تدريس OR د*رس OR اكتساب OR تعل*م OR تدريب OR تلقين OR بيداغوجي* OR بر*امج OR من*هج OR تل*ميذ OR ط*ل*ب OR دارس OR معلم OR اس*ت*ذ OR مدرس)	AND (“صوتي”OR “فك التشفير الصوتي” OR “إعادة الترميز الصوتي”OR “الترميز الصوتي” OR “فك التشفير الصوتي” OR “الفونولوجي”)	AND (مراجعة OR "مراجعة منهجية" OR استقصاء OR "تحليل كلي" OR "دراسة تجميعية")

3.5.Citation chaining

After finalizing the selection of eligible reports found through electronic searches, we reviewed the reference sections of the included papers to identify any additional potentially eligible reports that had not been previously identified.

3.6. Selection Process

After eliminating duplicate entries via Rayyan, all records underwent a preliminary screening based on their titles and abstracts to check for compliance with the set inclusion criteria. Records that failed to meet any one of our inclusion criteria were excluded, while those that could not be excluded moved forward to the full-text review phase. At the beginning of the process, a pair of independent reviewers independently evaluated a random 10% sample of the titles and abstracts, a procedure guided by the methodology of (Huang and Chalmers, 2023). Since the level of agreement between reviewers exceeded 90%, as suggested by Ramezanzadeh and Woore (2023), the screening of the remaining titles and abstracts was completed by the first author. Disagreements regarding the eligibility of studies were settled through discussion, with the rates of agreement duly noted in the review documentation. The next phase involved obtaining and closely examining the full texts of articles flagged as potentially relevant. In this phase, the methods section was read in detail, then other parts of the manuscript as necessary

to determine whether the paper met all inclusion criteria. This step repeated the double screening method against the criteria (10% random sample double screened. After 90% agreement, the remainder was screened by the first author). Articles failing to meet any one of the inclusion criteria at this stage were excluded, with the reasons for exclusion noted.

3.7.Data Extraction and Data Items

Records of titles and abstracts retrieved from multiple databases were uploaded into Rayyan, which is a web-based tool designed for joint abstract screening during the development of systematic reviews (Ouzzani *et al.*, 2016). This application incorporates advanced active machine learning strategies to expedite the screening process, and it has been shown to significantly enhance efficiency without detracting from the screening quality (Schoot *et al.*, 2021). For the data extraction process, Excel spreadsheets were used. For each article meeting inclusion criteria, study characteristics and data were extracted into a template in an Excel spreadsheet. An Excel spreadsheet served as the data extraction form for gathering information from the studies that met the eligibility criteria. A sample of the data extraction form is available in Appendix 1. The following data items were extracted from eligible studies: 1) Bibliographic information: Reference citation, year of publication, database/information source, publication type, language of publication, source of funding, 2) Setting of study: Context of study – i.e. country – (anglophone, non-anglophone), 3) Phonics instructions form: synthetic analytic and systematic or ad hoc, comparator, 4) Methods: Research questions, study design, measurement tools, hedges, 5) Participants: age (kindergarten, primary, high school, adult), setting (formal/informal), L1 (e.g., Arabic, Chinese), Orthographic (Roman/Non-Roman) 6) Outcomes: decoding skills, fluency, reading comprehension. The documentation of results encompassed statistical significance tests and measures of effect size. For cases where effect sizes are not specified in the original research, we calculated these, when necessary data were accessible, and 7) Limitations (author acknowledged; observed by researchers).

3.8.Quality Assessment

The credibility of each study was assessed using Gorard's (2014) framework. This methodology was specifically designed to evaluate the reliability of outcomes in experimental/intervention research. Regarding the full-text screening process, two separate reviewers independently assessed the risk of bias in 30% of the studies to ensure inter-rater reliability. We employed Gorard's Sieve (2015), a quality appraisal tool recommended for

educational research, to assess the trustworthiness of each study concerning design, scale, dropout rates, outcomes, fidelity, and validity. Each study received an overall star-rating ranging from 4*, indicating the most trustworthy evidence, to 0, indicating that the study did not adequately address sources of bias, if at all. According to the tool's guidelines, "an evaluation will be judged to be as good as the lowest classification it has achieved for each of the six categories" (Gorard, 2015: 6). *Design* was evaluated based on the comparability of groups at baseline, *scale* was assessed by sample size per group, and *dropout* was rated according to the level of attrition. *Outcomes* were judged based on the validity of the measurement methods, such as the use of standardized tests or validated researcher-designed tools. *Fidelity* was assessed by the clarity of the intervention and its delivery, while validity was rated on the representativeness of the results for similar EFL learners not involved in the studies. In cases where relevant information was missing, the item was rated as 0, following the tool's guidelines.

3.9.Data Synthesis

The results were organized based on criteria that emerged from the literature such as the type of phonics instruction (for instance, synthetic versus analytic, systematic versus ad hoc, length of instruction); the outcomes assessed (decoding/ fluency/ reading comprehension); characteristics of the learners (such as age, and whether their first language uses a Roman or non-Roman alphabet); and the setting (for example, anglophone – English speaking countries like the US or the UK; non-anglophone – non-English speaking countries like China). Initially, a numerical summary and a table detailed the characteristics and design aspects of the studies. Subsequently, a separate table showcased the thematic analysis of the findings from these studies. Since sufficient statistical data were unavailable and studies were not similar to support meta-analyses, a narrative synthesis was conducted instead, considering the reliability of the included studies. A narrative synthesis aims to produce "a summary of the current state of knowledge" in relation to the review question (Popay *et al.*, 2006: 6). This approach allowed for a comparison of findings from studies of varying quality, with more credibility given to those deemed more reliable. Where a body of literature includes diverse interventions and outcomes, as is often the case in social sciences research, it is inappropriate to conduct a meta-analysis of the studies' results (Petticrew and Roberts, 2008). Thus, following Petticrew and Roberts (2008), we conducted a narrative synthesis as follows: (i) studies are grouped into comparable categories based on outcome measures; (ii) findings and quality appraisal of

studies within each category are analysed; (iii) findings from all groups are synthesised narratively.

4. Results

4.1.Study Selection

After removing duplicates, the titles and abstracts of 3,264 records were screened. A total of 3,220 records were excluded based on the information provided. This resulted in 138 potentially eligible records, all of which were retrieved for full-text screening. Citation chaining identified 2 potentially eligible papers. After this screening, 25 studies were deemed to have met all the eligibility criteria and were included in the synthesis (see PRISMA flow diagram in Figure 1).

4.2.Study Characteristics

Table 3 provides characteristics for the 25 included studies. The themes emerging from the data are presented in a narrative synthesis.

4.3.Geographic context

Studies included in this review were conducted in 15 countries (Figure 2). All reviewed studies were in English as the Arabic studies did not meet the inclusion criteria.

4.4. Instructional context

Figure 3 illustrates the breakdown of participating settings. The majority of studies (n = 16) took place in primary schools, indicating a high awareness of the need for phonics instructions at that educational level. The remaining studies are split between preschool, university and informal institutes.

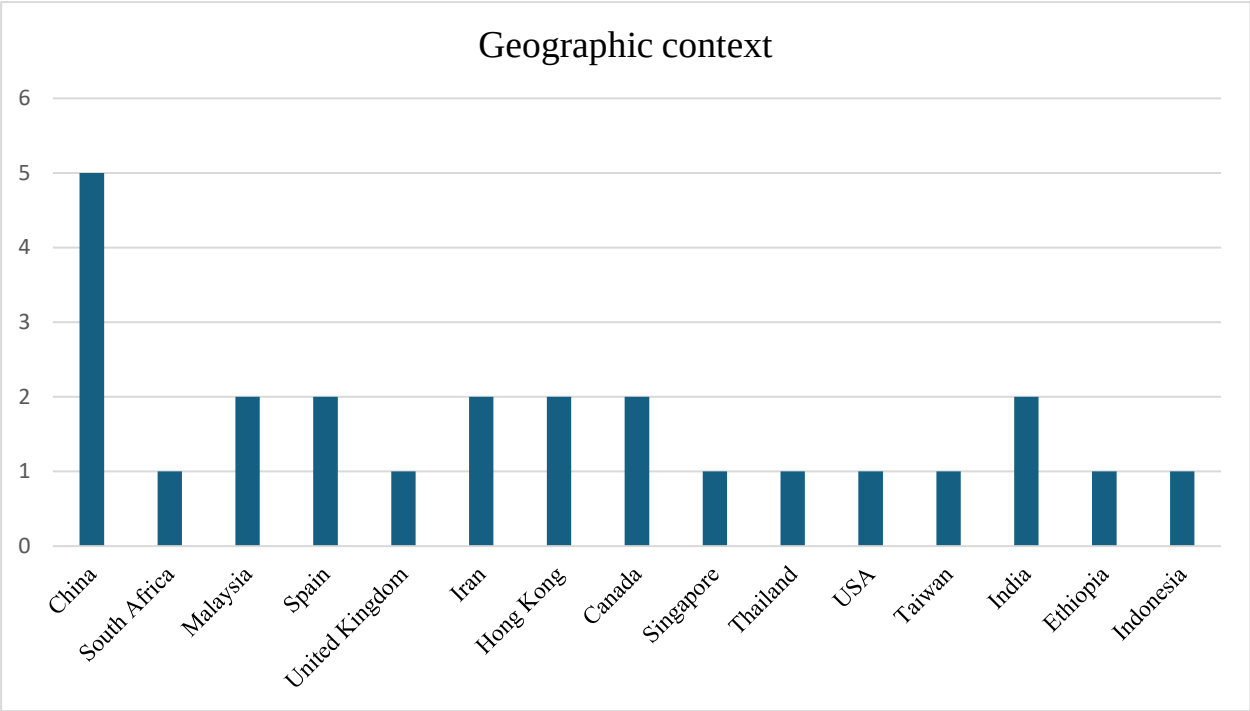

Figure 2 Frequency of geographic context of the studies

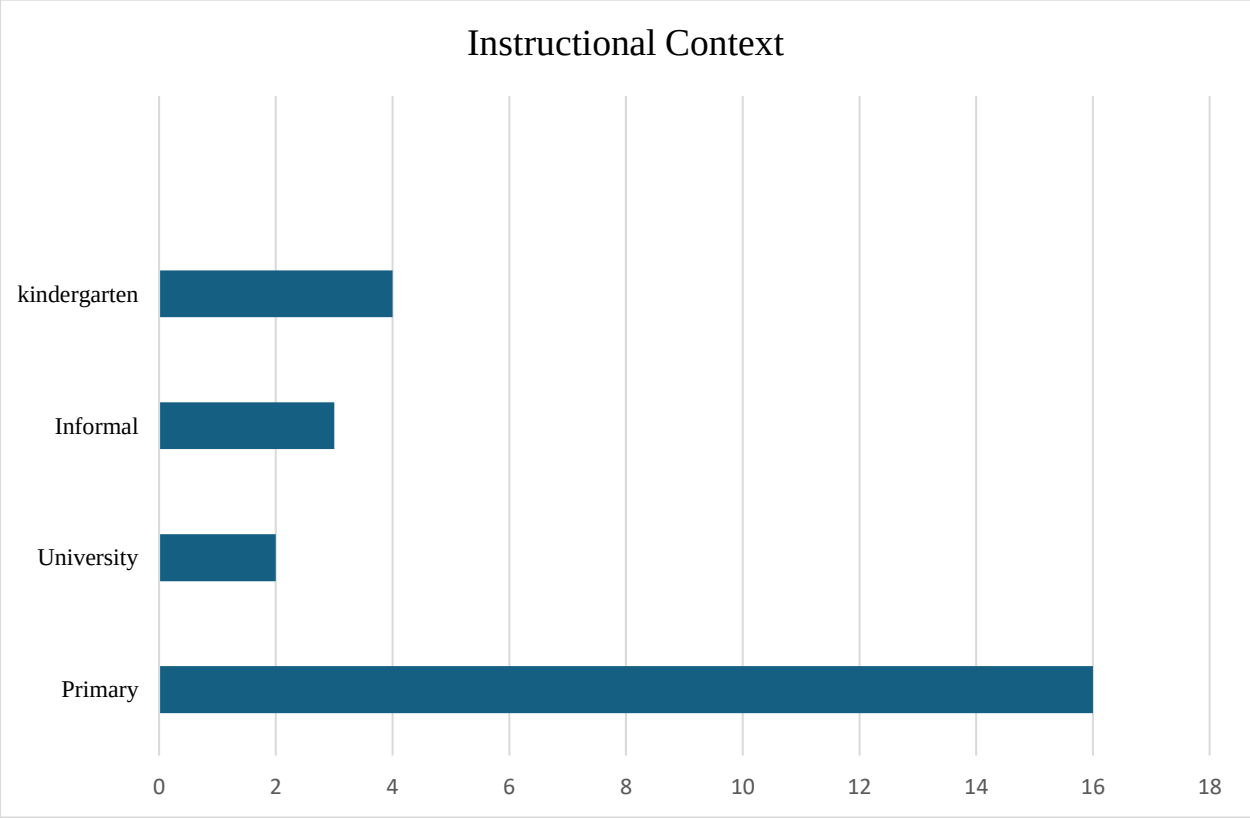

Figure 3 Frequency of instructional context of studies.

*Table 3 Study characteristics: study duration (*in weeks unless stated otherwise).*

Study	Country	Learner's L1	Study Design	Sample size	Setting	Study Duration (*weeks)	Specific Outcome Measures	Results	Quality Rating **
Li and Woore (2023)	China	Chinese	Quasi-experimental	138	University	12	English decoding accuracy (Woodcock Reading Mastery Test)	Improvement in decoding Pinyin-absent and some Pinyin-incongruent graphemes; no significant change in Pinyin-congruent graphemes.	3⋆
Le Roux *et al.* (2017)	South Africa	Setswana	Quasi-experimental	42	Primary	12	Test of Auditory Processing Skills (TAPS-3). One-Minute Reading Test. UCT Spelling Test.	Improved literacy and phonemic awareness in intervention group; significant differences in word discrimination and phonological segmentation.	2⋆
Rahman *et al.* (2023)	Malaysia	Malay	Quasi-experimental	60	preschool	4	Word-sound recognition and reading comprehension	Significant improvement in early reading proficiency.	2⋆
Li and Woore (2021)	China	Chinese	Quasi-experimental	138	University	12	Phonological decoding accuracy, decoding speed, vocabulary memorization (oral recall, written recall, aural recognition, written recognition)	Significant improvement in phonological decoding accuracy. Greater progress in recalling and recognizing spoken word forms and written recall of newly learned words.	2⋆
Li *et al.* (2017)	China	Chinese	pre-test, post-test, and follow-up assessments	112	Primary	3	English word recognition, pseudoword reading, phonological awareness (onset-rime and phonemic), Chinese character recognition, rapid naming (English letters and Chinese digits), and phonological memory (in both English and Chinese)	Poor English readers, especially those without Chinese-reading difficulties, significantly improved in English phonological awareness and word reading, with some sustaining these gains a year later.	2⋆
Fonseca-Mora *et al.* (2015)	Spain	Spanish	Quasi-experimental with pre-post comparison	63	primary	11	Wechsler Intelligence Scale for Children (WISC-IV), Early Grade Reading Assessment (EGRA), Musicality test.	Both groups (musical and non-musical) showed significant improvement in reading skills. No significant differences were found in terms of reading outcomes.	2★
Stuart (1999)	United Kingdom	86% of the children were English as a second language (ESL) learners, with the majority having Sylheti as their first language	Quasi-experimental, longitudinal study	112	Primary	12	Phoneme Awareness, Phonics Knowledge, Reading, and Writing.	The phonics group showed significant and lasting improvements in phoneme awareness, phonics knowledge, reading, and writing. One year after, they maintained their advantage, performing significantly better in reading and spelling tasks.	2★

Karimkhanlooei & Seifiniya (2015)	Iran	Turkish	Quasi-experimental	40	English Institute (3-6 yrs)	One year	Oral tests based on Sue Lloyd's phonics tests, assessing alphabet knowledge, reading, and writing skills	The phonics method group outperformed the traditional method group in overall scores, reading, and writing, with higher mean scores and lower standard deviations, indicating more consistent performance.	2★
Law *et al.* (2016)	Hong Kong	Cantonese	Quasi-experimental	30	Kindergarten	9	Phonological Assessment Battery (PhAB2 Primary): Alliteration Test and Rhyme Test.	The (computer-assisted intervention: CAI) group showed significant improvement in phonological processing skills compared to the control traditional group, as evidenced by higher scores in both the alliteration and rhyme tests immediately after the intervention and in the retention test 10 weeks later.	3★
Rendón-Romero *et al.* (2021)	Spain	Spanish	experimental pre-test – post-test	52	Primary	7 months	Spanish ALE1 Test, English Reading Predictors Test, Final Phonics Assessment.	The treatment group showed significant improvement in both Spanish and English literacy skills, particularly in phonological awareness and reading, compared to the control group.	3★
Lovett *et al.* (2008)	Canada	Portuguese: 37 Spanish: 16 Tagalog: 6 Italian: 3 Polish: 3 Arabic: 2 Syrian: 1 Urdu: 1 Unspecified: 7	Quasi-experimental, longitudinal	166	Primary	4 years	PPVT-3, CELF-3, CTOPP, WRMT-R Word Identification, Word Attack, Passage Comprehension, WRAT-3 Reading, Keyword Test, Sound–Symbol Test, Test of Transfer, Challenge Words Test	Significant improvements in reading and phonological processing skills for intervention group; no significant differences in outcomes between EFL and ELL students.	3★
O'Brien *et al.* (2022)	Singapore	Chinese	Randomized controlled trial	90	kindergarten	14	Baseline: Nonverbal reasoning, Receptive vocabulary (English, Mandarin), Word reading; Pretest/Posttest: Phonological awareness, Rapid symbol naming, Word reading, Decoding, Spelling; Follow-up: Same literacy measures	Greater growth in word reading, decoding, and spelling in the experimental group; Medium effect sizes for decoding and word reading, with sustained effect at follow-up; Certain grapheme-phoneme pairs were more challenging to learn.	**3★**
Limsukhawat *et al.* (2016)	Thailand	Thai	Quasi-experimental	36	primary	1.5 periods (regular English class)	Pretest and posttest scores on phonics understanding, Questionnaire on students' attitude towards the application	Significant improvement in students' phonics learning performance after using the application; positive attitudes towards the application with high satisfaction scores.	2★
Lesaux & Siegel (2003)	Canada	Predominantly Cantonese, Mandarin, Korean, Spanish, Persian, Polish, Farsi	Longitudinal, quasi-experimental	978 (790 L1, 188 ESL)	Primary	Kindergarten to Grade 2	Reading, spelling, phonological processing, memory, syntactic awareness, rapid naming	ESL students caught up to or outperformed L1 peers in Grade 2 on most reading measures; ESL students initially performed poorer on phonological processing and memory tasks but improved over time.	3★
Vadasy & Sanders (2012)	USA	Spanish, Vietnamese, Chinese, Somali, Amharic, Arabic, French, Russian, Samoan,	Quasi-experimental, longitudinal	187	Primary	20	Word reading (WRMT-R), Spelling (WRAT-IV), Reading comprehension (GORT-IV)	Significant long-term treatment effects on word reading, spelling, and reading comprehension, though smaller effects for English learners; no significant interaction with pretest word reading.	3★

		Tagalog, Cambodian							
Chu & Chen (2014)	Taiwan	Taiwanese (Mandarin)	Quasi-experimental	117	primary	5	English word reading (immediate and delayed post-tests)	Both groups improved in immediate post-tests, but the Phonics+ group (the phonics and the phonics +decodable text reading) showed significantly better retention in delayed post-tests, suggesting that decodable text reading enhanced long-term word reading retention.	3★
Patel *et al.* (2021)	India	Hindi	Randomized controlled trial	136	Primary	5	In-game measures: Letter-sound knowledge, Rime unit recognition, Word recognition; Out-of-game measures: Oral and paper-based English reading tasks	The phonics computer-assissted group showed significant improvements in in-game measures, but no significant differences were found on out-of-game measures; greater benefits observed in students with better pre-existing English literacy skills.	3★
Yeung *et al.* (2013)	Hong Kong	Cantonese	Quasi-experimental	76	Kindergarten	12	Phonological awareness (syllable deletion, rhyme detection, phoneme identification, etc.), English word reading, spelling, vocabulary	Phonics group showed significantly better performance in phonological awareness, word reading, spelling, and expressive vocabulary compared to the comparison group.	3★
Agegnehu *et al.* (2023)	Ethiopia	Amharic, Sidamuu Afoo	Quasi-experimental	70	primary	12	Phonological awareness (sentence segmentation, rhyme recognition and production, onset-rime segmentation and blending, phoneme segmentation, blending, deletion, substitution)	The experimental group showed significant improvement in phonological awareness across all subtests compared to the control group, demonstrating the effectiveness of the rime-based phonics method.	3★
Guo *et al.* (2023)	China	Chinese (Mandarin)	Quasi-experimental	339	primary	1 year	Phonological awareness, phonemic awareness, early literacy skills, initial letter sound fluency, segmenting, non-word reading	The experimental ICT phonics group group significantly outperformed the traditional phonics control group in all reading outcomes, with large effect sizes, demonstrating the effectiveness of ABRA in improving literacy skills.	3★
Tang and Peng (2019)	China	Chinese (Mandarin)	Quasi-experimental	6 (grade 1)	English training school	4 months	Listening, speaking, reading, writing, viewing	The translanguaging experimental group showed better comprehension and reduced confusion between Chinese and English phonics, performing significantly better in post-tests compared to the control group.	2★
Wahyuni (2023)	Indonesia	Indonesian	Pretest-posttest experimental	30	Primary	6 months	Word pronunciation, text comprehension, vocabulary mastery, confidence, and lesson enjoyment	Significant improvements in pronunciation, comprehension, vocabulary, and confidence; students and teachers reported positive experiences, with challenges noted in integrating phonics into regular curriculum.	3★
Jamaludin *et al.* (2016)	Malaysia	Malay	Quasi-experimental	80	Primary	14	Decoding, comprehension (based on adapted Phonological Awareness Literacy Screening and comprehension test)	The experimental group showed significantly higher post-test scores in decoding and comprehension compared to the control group, indicating the effectiveness of synthetic phonics instruction.	3★

Ahmadbeigi and Moloudi (2018)	Iran	Farsi	Quasi-experimental, mixed-methods	150	English Institutes	2 months	Reading accuracy, reading fluency, letter formation, spelling accuracy	Magic Phonics (locally customised phonics curriculum) users outperformed Jolly Phonics users in reading accuracy, fluency, letter formation, and spelling accuracy, particularly in the final phases of the study.	3★
Dixon *et al.* (2011)	India	Urdu, Telugu, Hindi, Arabic	Quasi-experimental	506	low-cost private unaided schools in Hyderabad	6 months	Burt reading test, Schonell spelling test, NFER-Nelson Diagnostic Reading Programme (letter matching, sound values, blending), dictation	Synthetic phonics intervention group showed significantly higher improvements in reading, spelling, letter recognition, and dictation compared to the traditional control group.	3★

4.5.Efficacy of Systematic Phonics Instruction

Systematic phonics instruction emerges as a highly effective method for improving literacy skills, particularly for ESL and EFL learners. Several studies have explored the impact of this approach. For instance, Li and Woore (2023) conducted a quasi-experimental study on Chinese EFL learners, implementing systematic phonics instruction that covered 101 grapheme-phoneme correspondences and 27 common word endings. Their findings demonstrated significant improvements in decoding unfamiliar English words and vocabulary learning, with the intervention group outperforming the comparison group, which only received phonological instruction without explicit phonics. Similarly, Vadasy and Sanders (2012) conducted a longitudinal study that followed lower-skilled first graders, including both English Learners (EL) and native English speakers, for two years after a systematic phonics intervention. Their study showed that the positive effects of phonics instruction on reading and spelling were sustained over time, particularly when reinforced by continued literacy instruction in the classroom. The study by Lesaux and Siegel (2003) demonstrates the crucial role of phonics instruction in helping EFL learners catch up with or surpass their native-speaking peers in reading proficiency by Grade 2. Initially, EFL students struggled with phonological processing and memory tasks, but consistent phonics instruction enabled significant improvement over time. This approach provides EFL learners with the essential skills to decode and understand English, bridging the gap between their first language and English phonology. The study underscores that phonics instruction is vital for long-term reading success among EFL learners. Lovett et al. (2008) examined the response to phonologically-based remediation programs among EFL struggling readers from different primary language backgrounds in a study involving 166 students. The study used a variety of phonological processing and reading assessments, including the CTOPP Blending Words subtest and the WRMT-R Word Identification subtest, to measure outcomes. The results indicated that while all students benefited from the intervention, EFL Learners showed more significant improvements in word identification and reading comprehension compared to their English as a First Language (EFL) peers. This research underscores the importance of targeted phonics instruction for EFL students. There seems to be a consistent positive effect for phonics instructions in the field of EFL. This body of research underscores the long-term benefits of systematic phonics for foundational literacy development.

4.6.Phonics Instruction in Multilingual and Diverse Contexts

Phonics instruction's effectiveness in multilingual and diverse settings is another key theme. Studies in this area focus on adapting phonics teaching to cater to students from different linguistic backgrounds. Le Roux et al. (2017) explored this in a quasi-experimental study involving Grade 3 Setswana L1 learners in South Africa. They implemented an intervention combining vowel perception and production techniques with phonological awareness activities, which resulted in significant improvements in the learners' phonemic awareness and literacy skills. The study highlighted the importance of tailoring phonics instruction to the specific phonological challenges faced by learners from different linguistic backgrounds. Similarly, Yeung, Siegel, and Chan (2013) examined the impact of a phonological awareness program on young Chinese ESL learners in Hong Kong. The study used a series of phonological awareness tasks, such as rhyme detection and phoneme identification, and found substantial improvements in the children's reading and spelling abilities. These findings suggest that when adapted to address the unique needs of learners from diverse linguistic backgrounds, phonics instruction can be highly effective.

4.7.Use of Technology and Innovative Approaches in Phonics Instruction

The integration of technology into phonics instruction is an emerging trend that enhances engagement and learning outcomes. Most previously conducted reviews either were conducted at a time frame before the COVID-19 pandemic, where TechEd emerged and the role of technology was emphasized, or had a strictly academic focus to assess the role of phonics instructions without looking at the factor of technology and therefore excluded all studies and interventions that had a technological aspect in it – whether be it an app – software or a website. This review explored the role of phonics instructions without overlooking the technological delivery methods used in the studies. O'Brien, Seward, and Zhang (2022) investigated the use of a multisensory interactive digital text, See Word Reading®, in a randomized controlled study with bilingual beginning readers in Singapore. The digital tool, which incorporated interactive activities on iPads, was found to significantly improve decoding, word reading, and spelling skills compared to traditional classroom phonics methods. The use of technology in this context provided an engaging and interactive platform that supported more personalized learning experiences. Another innovative approach was examined by Limsukhawat *et al.* (2016), who explored the effectiveness of synthetic phonics instruction delivered through an augmented reality-supported mobile game application. Their experimental study involved

struggling young ESL readers in Thailand and found that the technology-enhanced phonics instruction significantly improved early reading skills, suggesting that digital tools can offer effective support for literacy development in under-resourced settings.

The two studies by Patel et al. (2021) and Guo et al. (2023) both explored the use of technology in phonics instruction, but with differing results and implications for the role of technology in literacy education. Patel et al. (2021) examined the effects of the GraphoLearn (GL) English Rime program, a computer-assisted reading intervention, compared to a control group that used a math game. While the GL group showed significant improvements in in-game measures, such as letter-sound knowledge and word recognition, these gains did not translate to out-of-game measures, like oral and paper-based reading tasks. Moreover, the study found that students with better pre-existing English literacy skills benefited more from the GL program, suggesting that the intervention may not have been as effective for students with lower baseline skills. This outcome raises questions about the generalizability and practical impact of the technology-based intervention, particularly its ability to enhance real-world reading skills beyond the context of the game.

In contrast, Guo et al. (2023) demonstrated the substantial effectiveness of the ABRACADABRA (ABRA) program, a game-based online literacy tool, in significantly improving literacy outcomes across all measured domains, including phonics, segmenting, and non-word reading. The study was conducted in a rural setting and involved the experimental group using ABRA in addition to their regular English classes, while the control group received traditional English instruction without any interactive ICT tools. The experimental group outperformed the control group in every literacy measure, underscoring the effectiveness of integrating technology into phonics instruction. This study highlights the potential of well-designed educational technology to support and enhance traditional literacy teaching methods, particularly in under-resourced settings. It is clear that the success of technology in education depends significantly on how it is implemented and the context in which it is used. Patel et al. (2021)'s findings suggest that while technology can be engaging and beneficial within a controlled, game-based environment, its effectiveness may diminish when students are required to transfer those skills to more traditional reading tasks. This limitation could stem from a lack of integration between the technological tool and broader literacy instruction, suggesting that simply introducing technology is not enough; it must be seamlessly integrated into the overall curriculum to support skill transfer. On the other hand, Guo et al. (2023) study

demonstrates that when technology is effectively integrated into the classroom, with proper teacher training and support, it can lead to significant improvements in literacy outcomes, even in challenging environments. Therefore, while technology offers promising tools for literacy education, its implementation must be thoughtfully planned and supported to achieve the desired educational outcomes.

Another emerging innovative approach is the use of translanguaging in delivering phonics instructions. The study Rendón-Romero, Navarro-Pablo and García-Jiménez (2021) highlights the growing significance of translanguaging as an innovative method for teaching phonics, particularly in bilingual contexts. In this study, the treatment group, which engaged in translanguaging practices, demonstrated substantial improvements in both Spanish and English literacy skills, with marked gains in phonological awareness and reading proficiency. The findings suggest that translanguaging, by allowing students to draw on their full linguistic repertoire, supports the transfer of literacy skills between languages, thereby enhancing overall language development. This approach not only elevated the English reading proficiency of the treatment group to levels comparable with native English-speaking children but also underscored the potential of translanguaging to bridge language barriers and foster deeper linguistic and cognitive engagement in multilingual learners. As noted by other researchers, such as García and Wei (2014), translanguaging leverages bilingual students' entire linguistic knowledge, offering a holistic and flexible framework for language instruction that aligns with the complex realities of bilingualism in educational settings. However, the study was limited by the small sample size, and the role of translanguaging in delivering phonics instructions in bilingual contexts needs to be further explored before making generalizations.

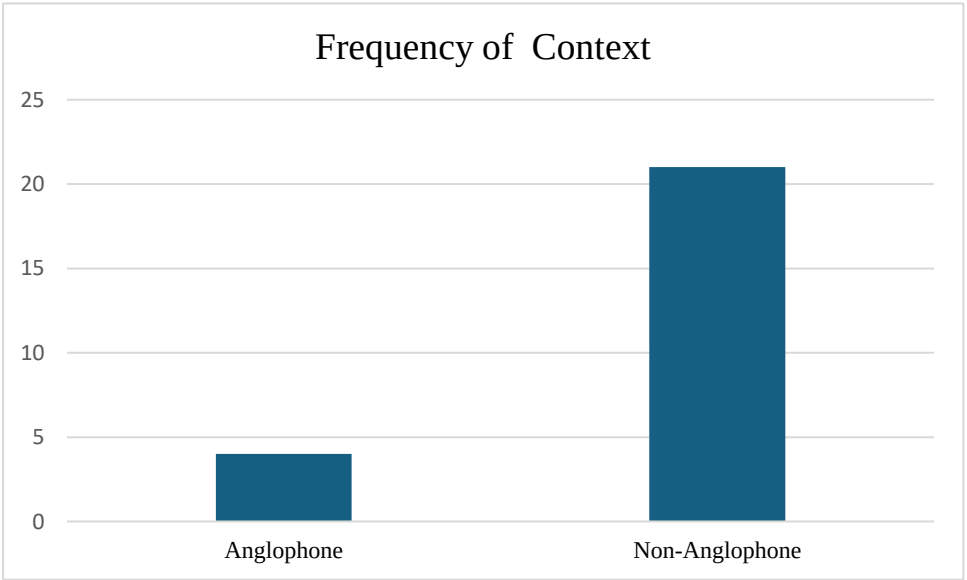

Figure 4 frequency of studies according to context..

4.8.Comparison of Effect Sizes Based on L1 and Context

The effectiveness of phonics instruction, as measured by effect sizes, varies significantly depending on the learners' L1 and whether the study was conducted in an anglophone (English-speaking countries) or non-anglophone (Non-English speaking countries) context. Figure 4 shows the frequency of studies according to context. Generally, studies in anglophone settings show moderate to large effect sizes, particularly when phonics instruction is provided to EFL learners within a supportive classroom environment. For instance, Lovett et al. (2008) reported large effect sizes (Cohen's d > 0.8) for improvements in word identification and phonological processing among EFL learners in Canada, suggesting that when phonics instruction is aligned with the linguistic needs of students, it can lead to substantial gains. Similarly, Vadasy and Sanders (2012) found significant long-term effects of phonics instruction with moderate to large effect sizes in their follow-up study of first-graders in the United States, particularly in reading fluency and comprehension.

In non-anglophone contexts, the effect sizes tend to be more variable. For example, Li and Woore (2023) reported a large effect size (r = .59) for phonological decoding accuracy among Chinese EFL learners, but this was contrasted by a smaller effect size in decoding speed (r = -.46), reflecting the challenges non-native speakers face when learning to read in English. Similarly, Le Roux et al. (2017) observed moderate effect sizes in phonemic awareness among Setswana-speaking students in South Africa, suggesting that while phonics instruction is effective, the complexity of transferring phonological skills across languages can temper the overall impact. Interestingly, studies that employed translanguaging approaches, such as the study by Tang and Peng (2019) in China, reported significant improvements across all skills measured, with moderate to large effect sizes (Cohen's d ranging from 0.69 to 1.12). This suggests that when the learners' L1 is actively integrated into instruction, it can enhance the effectiveness of phonics instruction, even in non-anglophone settings. This suggests that while phonics instruction is beneficial across different contexts, its effectiveness is often maximized in anglophone settings or when instruction is adapted to consider the learners' L1 and the linguistic environment. The strategic use of L1, as seen in translanguaging approaches, can also mitigate some of the challenges faced in non-anglophone contexts, leading to better literacy outcomes. When these findings are explored in the SVR model (Hoover & Gough, 1990), it reemphasizes the importance of developing listening skills in EFL learners, and that

phonics interventions in non-anglophone contexts should be tailored carefully around this consoderation.

4.9.Long-Term Impact and Sustainability of Phonics Instruction

The long-term impact of phonics instruction is a crucial consideration in evaluating its effectiveness. Vadasy and Sanders (2012) conducted a longitudinal follow-up study, assessing the effects of a kindergarten phonics intervention on literacy outcomes two years later. Their findings indicated that the benefits of phonics instruction were sustained, particularly among EFL Learners who received continued support through effective classroom literacy practices. The study used standardized assessments like the Woodcock Reading Mastery Test and Gray Oral Reading Test to measure outcomes such as word reading, spelling, and reading fluency, showing that students maintained or even improved their literacy skills over time. This research highlights the importance of sustained, high-quality phonics instruction in ensuring long-term literacy success.

4.10. Phonics Instruction vs. Other Literacy Approaches

The debate between phonics instruction and other literacy approaches, such as whole language or holistic methods, is another area of exploration. Stuart (1999) conducted a comparative study involving inner-city second language learners in London, where the effectiveness of phonics instruction was contrasted with a holistic reading approach using Big Books. The study found that the phonics group showed significantly greater gains in reading and spelling skills, as measured by tests of phoneme identification and segmentation, compared to the holistic group. Similarly, Wahyuni (2023) evaluated a synthetic phonics approach extended to storybook reading in a mixed-methods study with Indonesian third graders. The study used a combination of quantitative assessments, such as word reading and text comprehension tests, and qualitative data from teacher interviews. It found that integrating phonics with storybook reading significantly enhanced students' pronunciation, comprehension, and overall reading skills, suggesting that a phonics-blended approach can be highly effective.

4.11. Phonics in Early Childhood Education

The application of phonics instruction in early childhood education is critical for laying the foundation for literacy. Abdul Rahman, Tham, and Liu (2023) explored the use of early reading games based on phonics for preschoolers in Malaysia. Their quasi-experimental study involved 60 children, divided into intervention and control groups, with the intervention group receiving phonics-based early reading games. The study measured early reading proficiency using a series of diagnostics and found that the intervention group showed significant improvements in word-sound recognition and reading comprehension compared to the control group. Similarly, Karimkhanlooei and Seifiniya (2015) studied the effectiveness of phonics instruction for children aged 3-6 years in Iran, comparing it to traditional teaching methods. Their longitudinal study found that children taught using phonics showed superior alphabet knowledge, reading, and writing skills, as measured by custom assessments developed for the study, highlighting the effectiveness of phonics in early literacy development. Thus, the role of phonics instructions in early childhood should be more emphasized and explored.

4.12. Measures and Assessments Used in Phonics Studies

This review explored phonics studies that examined the role of phonics interventions on several measures. Figure 5 shows the frequency of studies according to each measure. In the field of literacy development, a significant body of research has focused on the impact of various instructional interventions on reading comprehension, particularly among young learners and those learning English as a second language. Across the reviewed studies, eight studies specifically reported on reading comprehension as a key outcome measure, offering valuable insights into the effectiveness of different approaches in diverse educational contexts (Lovett *et al.*, 2008; Vadasy & Sanders, 2012; Jamaludin *et al.*, 2016; Patel *et al.*, 2021; Rendón-Romero *et al.,* 2021; O'Brien *et al.,* 2022; Rahman *et al.,* 2023; Wahyuni, 2023). One notable study conducted in Malaysia, Jamaludin *et al.*, (2016) explored the effectiveness of synthetic phonics instruction among 80 primary school students. This quasi-experimental study measured outcomes in both decoding and reading comprehension. The intervention group, which received targeted phonics instruction, showed significantly higher post-test scores in reading comprehension compared to the control group. This finding aligns with the broader literature on phonics, which suggests that systematic phonics instruction can significantly enhance reading comprehension by improving decoding skills, thereby enabling students to access and understand text more effectively. Rahman, Tham and Liu (2023) in the Malaysian

context as well, examined the effects of an early literacy intervention on 60 Malay-speaking preschool children. This quasi-experimental study aimed to enhance word-sound recognition and reading comprehension, two foundational components of early reading proficiency. The results demonstrated significant improvements in reading comprehension within the intervention group, underscoring the importance of early and targeted interventions in preschool settings. The findings suggest that when literacy instruction is introduced at a young age, it can lead to marked advancements in comprehension, which is crucial for future academic success.

Similarly, Vadasy and Sanders (2012) conducted a longitudinal quasi-experimental study in the United States, involving 187 primary school students from various linguistic backgrounds, including Spanish, Vietnamese, and Chinese. The intervention focused on improving word reading, spelling, and reading comprehension. Throughout the study, the intervention group exhibited significant long-term gains in reading comprehension. However, it is worth noting that the effects were more pronounced among native English speakers compared to English learners. This outcome highlights the need for differentiated instructional strategies that address the unique challenges faced by EFL learners, ensuring that all students benefit equally from phonics interventions. These studies confirm the reliability of the SVR model, that phonics instructions can develop learners' ability to decode words but still need to understand spoken language to be proficient readers (Hoover and Gough, 1990). Thus, phonics interventions targeted at EFL learners should allow opportunities to develop listening comprehension for students to understand spoken language.

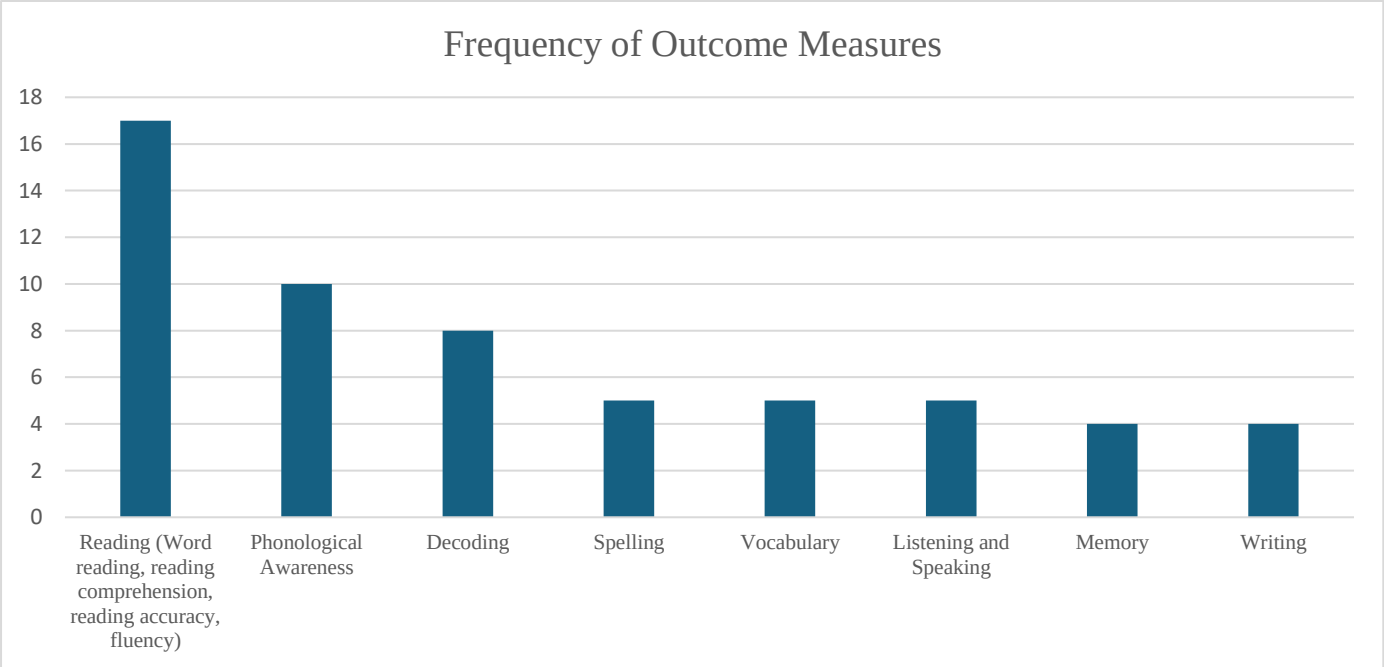

Figure 5 Frequency of studies according to each measure.

There is evidence in the literature that decoding is a fundamental skill in reading, enabling learners to translate written text into spoken words by accurately identifying and pronouncing letters and sounds (Ehri, 2005). This skill is particularly crucial for EFL learners, as it forms the basis for developing reading fluency and comprehension. Eight studies in this review explored the effect of phonics instructions on decoding measures (e.g. Jamaludin *et al.*, 2016; Li and Woore, 2021, 2023; O'Brien, Seward and Zhang, 2022). Studies such as those by Li and Woore (2023) and Chu and Chen (2014) emphasize the importance of decoding in the context of English language learning. Li and Woore (2023) found that when Chinese university students were challenged to decode English text without the aid of familiar Pinyin orthographic cues, their decoding accuracy and reading fluency significantly improved. This highlights the effectiveness of targeted phonics instruction that engages learners with the specific phonological challenges of a second language. Similarly, Chu and Chen (2014) demonstrated that incorporating decodable texts into phonics instruction not only improved decoding but also led to better retention of fluency skills over time. This approach allows learners to apply their decoding knowledge in meaningful reading contexts, reinforcing their ability to read accurately and fluently.

Further supporting the importance of decoding, Vadasy and Sanders (2012) conducted a longitudinal study showing that consistent phonics interventions significantly improved decoding and fluency among primary school students, particularly native English speakers. However, the study also highlighted the additional challenges faced by EFL learners, who, despite progress, did not achieve the same fluency gains as their native-speaking peers, indicating a need for more tailored support. Guo et al. (2023) expanded on this by exploring the role of technology in enhancing these skills, finding that ICT-based phonics instruction significantly outperformed traditional methods in improving decoding and fluency among Chinese primary school students. This underscores the potential of integrating technology into literacy education to meet diverse learning needs. Finally, Jamaludin *et al.* (2016) reinforced the effectiveness of synthetic phonics, which emphasizes systematic sound-symbol relationships, in developing decoding and fluency. The study's findings suggest that such structured phonics instruction enables students to decode words accurately and fluently, laying the groundwork for better reading comprehension and overall literacy (Adams, 1990). These studies demonstrate the critical role of decoding in reading development and the effectiveness of various instructional approaches in enhancing this skill across diverse educational contexts.

4.13. Effect Sizes and the Script Distance Hypothesis in Phonics Instruction

The effectiveness of phonics instruction can vary significantly depending on the learner's L1 script, particularly when comparing learners whose L1 uses a Roman alphabet with those whose L1 uses a non-Roman alphabet. Studies focusing on learners with Roman alphabet backgrounds, such as Spanish, Italian, or Portuguese speakers, often report larger effect sizes in phonological awareness, decoding, and reading fluency following phonics instruction. For instance, Lesaux and Siegel (2003) observed moderate to large effect sizes (Cohen's d = 0.60 to 1.24) in phonological processing and reading skills among Spanish-speaking English learners after targeted phonics interventions. Similarly, Fonseca-Mora, Jara-Jiménez and Gómez-Domínguez (2015)) documented significant improvements in phoneme awareness and reading fluency in young Spanish learners of English, with large effect sizes (Cohen's d = 1.04 for Phoneme Identification, Cohen's d = 0.92 for Phoneme Segmentation) when phonics instruction was supplemented with musical support. These findings suggest that learners from Roman alphabet backgrounds can more easily transfer their L1 literacy skills to English, thereby benefiting more substantially from phonics instruction.

In contrast, learners whose L1 uses a non-Roman script, such as Chinese, Arabic, or Korean, often exhibit smaller effect sizes in phonics-related outcomes. This is evident in studies such as that by Li and Woore (2021), where Chinese-speaking English as a Foreign Language (EFL) learners showed only moderate improvements in phonological decoding accuracy and vocabulary memorization (Cohen's d = 0.5-0.7). Another study by Guo et al. (2023) involving Chinese learners in rural China reported small to moderate effect sizes (Cohen's d = 0.25-0.55) for literacy improvement using a phonics-based digital tool. The comparatively smaller effect sizes in these studies reflect the greater difficulty that learners from non-Roman script backgrounds face in acquiring English phonics skills. These learners must not only learn a new set of grapheme-phoneme correspondences but also adapt to an entirely different orthographic system, making the process of acquiring phonological awareness and decoding skills more challenging. These differences can be understood through the lens of the script distance hypothesis, which posits that the greater the orthographic distance between a learner's L1 and the L2, the more difficult it will be for the learner to transfer existing literacy skills. Roman alphabet learners benefit from the similarity between their L1 and English, which facilitates the transfer of phonological awareness and decoding strategies. In contrast, learners from non-Roman script backgrounds experience a more pronounced script distance, which complicates

the acquisition of English literacy skills. This greater cognitive load can explain the smaller effect sizes observed in studies involving non-Roman script learners, as these individuals must navigate not only the linguistic differences but also the unfamiliar orthographic patterns of the English language. The script distance hypothesis provides a compelling framework for understanding the variability in phonics instruction outcomes based on learners' L1 scripts. The larger effect sizes observed in Roman alphabet learners indicate a smoother transfer of literacy skills to English, whereas the smaller effect sizes in non-Roman script learners underscore the additional challenges posed by orthographic differences. These findings suggest that phonics instruction may need to be tailored more carefully for non-Roman script learners to address the unique challenges they face in mastering English phonology and orthography.

5. Limitations

This review boasts several notable strengths, including a focused examination of studies of EFL learners, in-depth consideration of learners' characteristics, a recent scope of studies, and an emphasis on emerging trends in phonics instruction. However, it is essential to acknowledge the limitations of this review. Firstly, due to constraints in time and scope, this review did not explore grey literature or publications in languages other than English and Arabic. As a result, potentially valuable empirical evidence published outside mainstream academic sources or in other languages may have been overlooked, which could have contributed to a more comprehensive understanding of the subject matter. Secondly, the study adhered to strict inclusion criteria, which prioritized certain methodological designs over others. While such rigour is essential for maintaining the validity of the findings, it led to the exclusion of many qualitative and observational studies. Incorporating observational and qualitative studies into systematic reviews, particularly those assessing intervention effectiveness, presents significant challenges. These challenges include an increased risk of bias, greater heterogeneity, and difficulties in synthesizing diverse data. Observational studies are particularly prone to biases and confounding variables, which can skew results. Meanwhile, qualitative studies, despite their value in exploring context-specific factors, often lack the generalizable and quantifiable data necessary for robust meta-analyses. The exclusion of these studies was intended to preserve the validity and reliability of the review's conclusions. As supported by Chalmers (2021), the prioritization of randomized controlled trials is critical in systematic reviews to maintain rigour and ensure that findings are broadly applicable. Moreover, studies conducted in Arabic did not meet the inclusion criteria, further narrowing the scope of this review. While the original research protocol included an intention to evaluate the confidence in the aggregated evidence using the GRADE methodology, which involves a thorough assessment of bias, accuracy, consistency, relevance, and the risk of publication bias, this comprehensive evaluation could not be conducted due to time and resource limitations. It is anticipated that future updates to this review will address these limitations by incorporating a broader range of studies and employing the GRADE methodology to provide a more nuanced evaluation of the evidence quality.

6. Conclusions

The findings of this study confirm the effectiveness of phonics instruction in improving English literacy among EFL learners, with notable differences based on the learners' first language script and context. Learners from Roman script backgrounds benefit more significantly due to the easier transfer of phonological and orthographic skills, while those from non-Roman scripts face greater challenges, reflected in smaller effect sizes. These results highlight the importance of tailoring phonics instruction to address the specific needs of EFL learners. In particular, adapting phonics interventions for non-Roman script learners should not only focus on reading skills but also incorporate listening comprehension strategies to enhance overall language acquisition. By considering script distance and listening comprehension in instructional approaches, educators can better support learners in acquiring English literacy skills. Future research should focus on refining these tailored strategies to maximize the effectiveness of phonics instruction across diverse linguistic backgrounds. Additionally, emerging trends such as the integration of technology and translanguaging are promising delivery methods that could further enhance phonics instruction. However, these approaches need to be rigorously examined in various contexts and with larger sample sizes to fully understand their impact and effectiveness.

References

Adams, M.J. (1990) *Beginning to read: Thinking and learning about print*. MIT Press.

Ahmadbeigi, N. and Moloudi, M. (2018) 'Effect of locally customized vs. native speakers' phonics curriculum on young Iranian EFL learners' early literacy development', I*ndian Journal of Science Communication*, 17(1), pp. 3–18.

Albalooshi, E. (2022) *The Role of a Phonic Intervention Program in Developing Reading Skills in Second Language Acquisition of English for Struggling Readers in Bahrain*. The University of Sheffield.

Association, N.G. (2010) 'Common core state standards'. Washington, DC.

August, D. and Shanahan, T. (2006) 'Executive summary', in *Developing Literacy in Second-Language Learners: A Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

August, D., Shanahan, T. and Escamilla, K. (2009) 'English language learners: Developing literacy in second-language learners—Report of the National Literacy Panel on Language-Minority Children and Youth', *Journal of literacy research*, 41(4), pp. 432–452.

Bassetti, B. (2006) 'Orthographic input and second language phonology: The influence of the L2 orthography on L2 phonological awareness', *Language and Cognitive Processes*, 21(5), pp. 473–500.

Catts, H.W., Adolf, S.M. and Weismer, S.E. (2006) 'Language deficits in poor comprehenders: A case for the simple view of reading', *Journal of Speech, Language, and Hearing Research*, 49, pp. 278–293.

Cenoz, J. and Gorter, D. (2015) *Multilingual education: Between language learning and translanguaging*. Cambridge University Press.

Dilgard, C., Hodges, T.S. and Coleman, J. (2022) 'Phonics Instruction in Early Literacy: Examining Professional Learning, Instructional Resources, and Intervention Intensity', *Reading Psychology*, 43(8), pp. 541–575. Available at: https://doi.org/10.1080/02702711.2022.2126045.

Dixon, P., Schagen, I. and Seedhouse, P. (2011) 'The impact of an intervention on children's reading and spelling ability in low-income schools in India', *School Effectiveness and School Improvement*, 22(4), pp. 461–482.

Education Science, D. and Training (2005) *National Inquiry into the Teaching of Literacy (NITL*. Teaching reading: Report and recommendation.

Ehri, L.C. (2005) 'Learning to read words: Theory, findings, and issues', *Scientific Studies of Reading*, 9(2), pp. 167–188.

García, O. (2009) *Bilingual Education in the 21st Century: A Global Perspective*. Wiley-Blackwell.

García, O. *et al.* (2017) *The translanguaging classroom: Leveraging student bilingualism for learning*. Philadelphia, PA: Caslon.

García, O. and Wei, L. (2014) *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan.

Gersten, R. *et al.* (2007) in *Effective Literacy and English Language Instruction for English Learners in the Elementary Grades. IES Practice Guide. NCEE 2007-4011*. What Works Clearinghouse.

Gorard, S. (2014) 'The widespread abuse of statistics by researchers: What is the problem and what is the ethical way forward?', *The Psychology of Education Review: the journal of the Education Section of the British Psychological Society*, 38(1).

Gorard, S. (2021) 'How to make sense of statistics'.

Gorard, S.A.C. (2015) 'Rethinking 'quantitative methods and the development of new researchers', *Review of Education*, 3(1).

Gough, P.B. and Tunmer, W.E. (1986) 'Decoding, Reading, and Reading Disability', *Remedial and Special Education*, 7(1), pp. 6–10. Available at: https://doi.org/10.1177/074193258600700104.

Hall, A., Higgins, S. and Wall, K. (2019) 'Technology-enhanced learning and phonics: An exploration of current practices and future directions', *British Journal of Educational Technology*, 50(4), pp. 1743–1757.

Higgins, S. and Xiao, Z. (2020) 'Integrating technology into phonics instruction: Challenges and opportunities', *Computers & Education*, 153, p. 103900.

Hindman, A.H. *et al.* (2020) 'Bringing the science of reading to preservice elementary teachers: Tools that bridge research and practice', *Reading Research Quarterly*, 55, pp. 197–206.

Hoover, W.A. and Gough, P.B. (1990) 'The simple view of reading', *Reading and writing*, 2, pp. 127–160.

Huang, X. and Chalmers, H. (2023) 'Implementation and effects of pedagogical translanguaging in EFL classrooms: A systematic review', *Languages*, 8(3), p. 194.

Jamaludin, K.A. *et al.* (2016) 'The effectiveness of synthetic phonics in the development of early reading skills among struggling young ESL readers', *School Effectiveness and School Improvement*, 27(3), pp. 455–470.

Koda, K. (2005) *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Le Roux, M. *et al.* (2017) 'Phonemic awareness of English second language learners', *South African Journal of Communication Disorders*, 64(1), pp. 1–9.
Li, S. and Woore, R. (2021) 'The effects of phonics instruction on L2 phonological decoding and vocabulary learning: an experimental study of Chinese EFL learners', *System*, 103. Available at: https://doi.org/10.1016/j.system.2021.102677.
Li, S. and Woore, R. (2023) 'How Chinese Learners Decode L2 English Words: Evidence from a Phonics Instruction Program'.
Lipka, O. and Siegel, L.S. (2007) 'The development of reading skills in children with English as a second language', *Scientific Studies of Reading*, 11(2), pp. 105–131.
Lipka, O., Siegel, L.S. and Vukovic, R. (2005) 'The literacy skills of English language learners in Canada', *Learning Disabilities Research & Practice*, 20(1), pp. 39–49.
McCarthy, A., Rickard, N. and Alexandra, S. (2018) 'The impact of digital tools on phonics instruction in early literacy education', *Journal of Educational Technology & Society*, 21(2), pp. 125–137.
Murphy Odo, D. (2021) 'A Meta-Analysis of the Effect of Phonological Awareness and/or Phonics Instruction on Word and Pseudo Word Reading of English as an L2', *SAGE Open* [Preprint].
Ouzzani, M. *et al.* (2016) 'Rayyan—a web and mobile app for systematic reviews', *Systematic Reviews*, 5, pp. 1–10.
Rahman, M.N.A., Tham, K.E. and Liu, C. (2023) 'Implementation of Early Reading Games to Teach English Using Integrated Sound Word Method for Preschoolers'.
Ramezanzadeh, A.M. and Woore, R. (2023) 'Ramezanzadeh and Woore (2023)-Scoping Review of Teaching and Learning Arabic as an L2-Database'.
Rendón-Romero, S.I., Navarro-Pablo, M. and García-Jiménez, E. (2021) 'Using phonics to develop the emergent English literacy skills of Spanish learners', *Porta Linguarum* [Preprint].
'Report of the National Reading Panel: Teaching children to read' (2000). U.S. Government Printing Office.
Rose, J. (2006) 'Independent review of the teaching of early reading: Final report'.
Shankweiler, D. *et al.* (1999) 'Comprehension and decoding: Patterns of association in children with reading disabilities', *Scientific Studies of Reading*, 3(1), pp. 69–94.
Sparks, R.L. (2021) 'Identification and Characteristics of Strong, Average, and Weak Foreign Language Readers: The Simple View of Reading Model', *The Modern Language Journal*, 105(2), pp. 507–525. Available at: https://doi.org/10.1111/modl.12711.
Spencer, M. and Wagner, R.K. (2017) 'The comprehension problems of children with poor reading comprehension despite adequate decoding: A meta-analysis', *Review of Educational Research*, 87(3), pp. 366–400.
Stanovich, K.E. (1993) 'The construct of phonological awareness', *Journal of Educational Psychology*, 85(1), pp. 1–20.
Stuart, M. (1999) 'Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners'.
Tharp, R.G. (1991) *The instructional conversation: Teaching and learning in social activity*. National Center for Research in Cultural Diversity and Second Language Learning.
Velasco, P. and García, O. (2014) 'Translanguaging and the Writing of Bilingual Learners', *Bilingual Research Journal*, 37(1), pp. 6–23.
Wang, M., Koda, K. and Perfetti, C.A. (2003) 'Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners', *Cognition*, 87(2), pp. 129–149.
Washburn, E.K., Joshi, R.M. and Binks Cantrell, E. (2011) 'Are preservice teachers prepared to teach struggling readers?', *Annals of dyslexia*, 61, pp. 21–43.
Woore, R. (2021) 'Teaching phonics in a second language', in *Debates in second language education*. Routledge, pp. 222–246.
Yeung, S.S., Siegel, L.S. and Chan, C.K. (2013) 'Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children', *Reading and writing*, 26(5), pp. 681–704.

Appendix A blank data extraction form

Category	Item	Description
Bibliographic information	Reference citation Year of publication Database/information source Publication type Language and location of publication Source of funding	Include bibliographic information about the study.
Setting of study	Context of study	Indicate the context and country of the study to determine whether it is in anglophone (English speaking country), or non-anglophone (non-English speaking country).
Intervention	Language of instruction Description of Phonics instructions Duration Comparison (if any) Number of participants Class grouping Baseline imbalances Attrition	Which language was used? What type of phonics instruction was used? (e.g. synthetic analytic and systematic or ad hoc). How long did the intervention last? Was there any control group included in the study? If so, how is the control group different from the experimental group? Include breakdown of the number of participants: n = number of participants n = intervention group n = control group How were the participants assigned to groups? Were there any notable differences between groups at the start of the study? Did any participants drop out of the study? If so, how many, and what were the reasons?
Methods	Research questions Study design	Indicate the research questions and design of the study
Participants	Age/gender Setting L1 (e.g., Arabic, Chinese) Orthographic	kindergarten, primary, high school, adult. Is the setting formal or informal? What is the L1, L2, L3 of the learners? (e.g. Roman/Non-Roman alphabet) Include any information regarding participants' learning disabilities, socioeconomic background, etc. what is the proficiency level at beginning of the study in each language? What is the age range and the number at each age? Include gender breakdowns where available.
Outcomes	Outcome type Outcome name Unit(s) of measure Time points measured Descriptive outcomes	Indicate the language skills measured (e.g. decoding skills, fluency, reading comprehension.) How is outcome operationalised? How many data collections? When?
Results	Statistical significance tests Measures of effect size	Indicate specific names of the measures and tests used. For cases where effect sizes are not specified in the original research, we calculated these, when necessary data were accessible.
Limitations	Indicate the limitations of the study	author acknowledged; observed by researchers