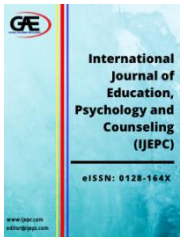
International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228



INTERNATIONAL JOURNAL OF EDUCATION, PSYCHOLOGY AND COUNSELLING (IJEPC)
https://gaexcellence.com/ijepc

A STRUCTURED LITERATURE REVIEW OF TEXT COMPLEXITY IN CHILDREN'S STORYBOOKS

Anealka Aziz [1]*, Muna Liyana Mohamad Tarmizi[2], Suzana Hamzah[3], Ng Yu Jin[4], Najihah Mohammad Yusoff [5], Mallika Vasugi V. Govindarajoo[6]

[1] Akademi Pengajian Bahasa, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
anealka@uitm.edu.my https://orcid.org/0000-0001-9280-242X
[2] Akademi Pengajian Bahasa, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
muna4566@uitm.edu.my https://orcid.org/0000-0002-9999-2614
[3] Akademi Pengajian Bahasa, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
suzana@uitm.edu.my https://orcid.org/0009-0000-2798-8969
[4] UNITAR University College, Kuala Lumpur, Malaysia
yujin.ng@unitar.my https://orcid.org/0000-0002-2993-4443
[5] UNITAR University College, Kuala Lumpur, Malaysia
najihah.yusof@unitar.my https://orcid.org/0009-0004-5454-6196
[6] UNITAR International University, Petaling Jaya, Selangor, Malaysia
mallika@unitar.my https://orcid.org/0009-0005-3222-9883
*Corresponding Author

Article Info:

Article history:
Received date: 30.09.2026
Revised date: 15.10.2026
Accepted date: 05.11.2026
Published date: 01.12.2026

To cite this document:

Aziz, A., Tarmizi, M. L. M., Hamzah, S., Ng, Y. J., Yusoff, N. M., & Govindarajoo, M. V. V. (2026). A Structured Literature Review of Text Complexity in Children's Storybooks. *International Journal of*

Abstract:

Text complexity matters when judging whether children's storybooks are suitable for young readers in terms of language, development, and meaning. Yet studies of text complexity often approach the issue from different angles. Some focus on measurement, others on linguistic features, and others on how children respond to the text. As a result, these strands of research are not always brought together. This study reviews research on text complexity in children's storybooks published from 2017 to 2026. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. Searches were conducted in Scopus and Web of Science using terms related to text complexity, readability, linguistic complexity, and children's literature. After screening, duplicate removal, eligibility checking, and quality appraisal, 15 studies were included in the qualitative synthesis. The analysis produced three main themes. The first, Computational Measures of Text Complexity and Readability, covers readability formulas, automated linguistic measures, corpus-based resources, and computational modelling. The second, Linguistic Dimensions of Text Complexity, includes lexical, morphological,

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

Education, Psychology and Counselling, 11(65), 207-228.

syntactic, grammatical, and discourse-related features. The third, Reader-Text Interaction and Text Complexity, considers children's language resources, shared reading, social and technological support, presentation mode, and the relationship between visual and verbal information. The findings show that readability scores alone are not enough to explain text complexity. Difficulty also depends on the linguistic features of the text and on the conditions in which children engage with it. Based on these findings, the review brings three perspectives together. Computational measures show how complexity can be assessed, linguistic analysis helps explain which features make a text more demanding, and reader-text research shows how those demands may be experienced by children. Taken together, these perspectives provide a broader basis for selecting, evaluating, adapting, and developing storybooks for young readers.

DOI: 10.35631/IJEPC.1165011

Keyword:

Children's Storybooks; Children's Literature; Readability; Linguistic Complexity; Text Complexity

© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact ijepc@gaexcllence.com.

Introduction

Storybooks for children play an important role in language acquisition, literacy development, and cognitive growth. Their suitability for young readers depends partly on text complexity, which influences comprehension, engagement, and learning. Traditionally, text difficulty has often been assessed through readability measures based primarily on word and sentence characteristics. However, these measures may not adequately capture semantic, syntactic, discourse, narrative, and multimodal demands that influence how children understand texts (Battistelli et al., 2022; Fitzgerald et al., 2015; Tortorelli, 2020).

Text complexity in children's storybooks is therefore better understood as a multidimensional construct. Vocabulary, syntactic structure, narrative organisation, thematic content, illustrations, and the developmental characteristics of readers may all influence how difficult or accessible a storybook is (Balaban & Hohenberger, 2020; Hsiao et al., 2024; Tortorelli, 2020). These dimensions may also interact; for example, a linguistically demanding text may be supported by illustrations or familiar narrative content, while apparently simple language may still require considerable inference from young readers.

Despite growing research on these dimensions, studies of children's text complexity remain distributed across different perspectives, including readability, linguistic analysis, narrative research, developmental studies, and multimodality. This fragmentation makes it difficult to determine how the field has conceptualised and assessed text complexity. A systematic synthesis is therefore needed to clarify how text complexity is measured,

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228
which linguistic features contribute to complexity, and how textual demands interact with readers and the conditions under which storybooks are encountered.

Literature Review

Text complexity in children's literature extends beyond conventional readability measures. Historically, textual difficulty has been evaluated using quantitative formulas based on surface-level properties such as word length, syllable counts, and sentence length (Günaydın, 2023; Imperial & Ong, 2021; Zhao et al., 2022). Although useful for estimating difficulty, such measures may produce inconsistent results and often overlook discourse-level characteristics such as cohesion and coherence, as well as the interaction between texts and readers (Alruthaya et al., 2025; Günaydın, 2023; Zhao et al., 2022). This has encouraged greater attention to linguistic, narrative, developmental, and multimodal dimensions of complexity.

At the linguistic level, lexical and syntactic characteristics are important indicators of storybook difficulty. Corpus-based studies have used measures such as lexical diversity, type-token ratio, word-frequency distributions, lexical density, and syntactic complexity to differentiate texts across reading or grade levels (Kupriyanov et al., 2023; Zhao et al., 2022). Zhao et al. (2022), for example, found vocabulary diversity, verbs per utterance, and morpheme length to be associated with expert judgements of picture-book difficulty. Computational studies also show that text complexity is better captured when lexical, syntactic, and part-of-speech features are considered together, rather than relying mainly on measures such as sentence or word length (Imperial & Ong, 2021; M. Liu et al., 2023). Other features, including complex nominal groups, clausal subordination, and argument structures, may also make texts more demanding for young readers (Andreeva et al., 2025; Hartati et al., 2025; Muhinyi et al., 2025).

Text complexity is not limited to words and sentences. The way a story is organised can also affect how easily readers follow events, characters, and ideas. Short sentences, for instance, do not always make a text easier if causal links are left unstated and readers have to infer them (Alruthaya et al., 2025; Kupriyanov et al., 2023). Comprehension may also be affected by multi-episode plots, character goals, the need to make inferences, and more abstract ideas (Kawar et al., 2023; Muhinyi et al., 2025; Schwarz et al., 2024). In other words, simple language does not always mean a simple story.

The multimodal features of children's storybooks add another layer of complexity. Illustrations, print, and layout can support understanding, but they may also make reading more demanding when the visual and verbal information do not match (Alruthaya et al., 2025). Children may be able to read the words correctly but still have difficulty linking them to information shown in the illustrations. The format of the storybook can also shape interaction during shared reading. Wordless books, for example, may encourage more discussion and adult support, while written texts may provide greater lexical and syntactic variety (Petrie et al., 2021). Visual and verbal information, therefore, should be considered together when examining text complexity.

Research also shows that children's reading materials do not always become steadily more complex with age or grade level. Longitudinal studies have found changes in the linguistic demands of elementary reading materials over time, while research on storybook series shows that progression from easier to more difficult texts can be uneven (Günaydın, 2023; Hiebert, 2025). Complexity also differs across genres. Informational texts, for example, often contain a

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228
wider range of vocabulary, more varied sentence structures, and denser information than simpler narratives (Andreeva et al., 2025; Gibbs & Reed, 2024). This variation suggests that storybook complexity is better examined through several features rather than a single readability score.

Research Questions

Formulating research questions (RQs) is an important stage in a Structured Literature Review, as the RQs establish the scope and direction of the review. Clearly defined research questions guide the identification, selection, analysis, and synthesis of relevant studies while keeping the review aligned with its objectives. They also provide a systematic structure for examining the selected literature and synthesising evidence across studies. Well-focused RQs help identify recurring patterns, methodological trends, key findings, and gaps in existing research more systematically. In addition, clearly articulated research questions improve the review's transparency and reproducibility by establishing a clear connection between the study's purpose, the literature search, and the subsequent analysis.

The formulation of research questions is therefore a critical component of the planning stage because it influences the subsequent stages of the review process (Kitchenham, 2007). In the present review, which examines text complexity in storybooks for children, the PICo framework proposed by Lockwood et al. (2015) was adopted to guide the development of the research questions. PICo consists of three components: Population (P), Interest (I), and Context (Co). For this review, the population refers to children or young readers, the phenomenon of interest concerns the different dimensions and assessment of text complexity, and the context comprises storybooks and children's literature. The framework provides a systematic way to define the review's boundaries while ensuring the research questions align with its central focus.

Based on the PICo framework and the objectives of the review, three research questions were formulated to guide the identification, analysis, and synthesis of the selected studies. The questions address three complementary aspects of text complexity: its computational measurement, the linguistic features that contribute to complexity, and the interaction between textual demands and readers. The following research questions guided the review:

- RQ1: What computational measures are used to assess text complexity and readability in storybooks for children?
- RQ2: What linguistic features are identified as contributing to text complexity in storybooks for children?
- RQ3: How does reader-text interaction shape the experience of text complexity in storybooks for children?

Materials and Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework by Page et al. (2021) was used in conducting the systematic literature review for this study. PRISMA was adopted to ensure transparency, rigour and replicability throughout the review process. It provided a systematic procedure to identify, screen, select and report the relevant studies. Two bibliographic databases of established

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

indexing standards, Scopus and Web of Science (WoS), were selected to source the articles for analysis. These two were chosen for their broad multidisciplinary coverage and peer-reviewed scholarly articles. The PRISMA framework consists of four main stages: Identification, Screening, Eligibility and Inclusion.

Identification

Using predefined keywords and Boolean operators, systematic searches were conducted in Scopus and WoS to identify studies relevant to text complexity in children's storybooks. This search strategy helps ensure adequate coverage and relevance to the review scope. The search included keywords *"text complexity," "children's storybooks," "readability," "linguistic complexity," and "children's literature."* These keywords were chosen to capture studies addressing both the linguistic characteristics of texts and their suitability for young readers, and to provide an appropriate starting point to achieve breadth without diverting from the scope of the review.

The database searches yielded 443 records before deduplication and screening. 240 records (54.2%) were retrieved from Scopus, while 203 (45.8%) were retrieved from WoS. These 443 records represented potentially relevant publications, not yet the final number for analysis. These records were then screened to remove duplicates and titles and abstracts that did not meet the predetermined inclusion and exclusion criteria. Table 1 shows the search strategies used in the study.

Table 1. Search Strategies Applied to Scopus and Web of Science

Scopus	TITLE-ABS-KEY (("text complexity" OR "linguistic complexity" OR "text difficulty" OR readability OR "lexical complexity" OR "lexical diversity" OR "lexical sophistication" OR "vocabulary complexity" OR "vocabulary difficulty" OR "syntactic complexity" OR "sentence complexity") AND (storybook* OR "story book*" OR "children's book*" OR "children's literature" OR "children's stor*" OR "picture book*" OR picturebook* OR "narrative text*" OR fable* OR folktale* OR "fairy tale*")) AND (LIMIT-TO (PUBYEAR , "2017") OR LIMIT-TO (PUBYEAR , "2018") OR LIMIT-TO (PUBYEAR , "2019") OR LIMIT-TO (PUBYEAR , "2020") OR LIMIT-TO (PUBYEAR , "2021") OR LIMIT-TO (PUBYEAR , "2022") OR LIMIT-TO (PUBYEAR , "2023") OR LIMIT-TO (PUBYEAR , "2024") OR LIMIT-TO (PUBYEAR , "2025") OR LIMIT-TO (PUBYEAR , "2026")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (OA , "all")) Date of Access: October 2026
WoS	Refine results for ("text complexity" OR "linguistic complexity" OR "text difficulty" OR readability OR "lexical complexity" OR "lexical diversity" OR "lexical sophistication" OR "vocabulary complexity" OR "vocabulary difficulty" OR "syntactic complexity" OR "sentence complexity") AND (storybook* OR "story book*" OR "children's book*" OR "children's literature" OR "children's stor*" OR "picture book*" OR picturebook* OR "narrative text*" OR fable* OR folktale* OR "fairy tale*") (Topic) and

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

	Open Access and 2026 or 2025 or 2024 or 2023 or 2022 or 2021 or 2017 or 2018 or 2019 or 2020 (Publication Years) and Open Access and 2026 or 2025 or 2024 or 2023 or 2022 or 2021 or 2018 or 2017 (Publication Years) and Article (Document Types) and Open Access and 2026 or 2025 or 2024 or 2023 or 2022 or 2021 or 2018 or 2017 (Publication Years) and Article (Document Types) and English (Languages) and All Open Access (Open Access) and Article (Document Types) Date of Access: October 2026

Screening

After identifying 443 records, they underwent a screening process based on pre-defined eligibility criteria. The study excluded records published before 2017, written in languages other than English, classified as conference proceedings, conference papers, books, reviews, or in-press publications, indexed outside the Social Sciences or Arts and Humanities subject areas, or unavailable through open access. 348 records (78.6%) were excluded, leaving 95 records for further assessment (56 Scopus and 39 WoS). 26 duplicated records from Scopus and WoS were identified and removed, leaving 69 unique records for further screening by title and abstract and eligibility assessment. This deduplication step prevented multiple representation of the same study and ensured each record contributed only once in the evidence-selection process. The inclusion and exclusion criteria are displayed in Table 2.

Table 2. Inclusion and Exclusion Criteria for Study Selection

Criterion	Inclusion Criteria	Exclusion Criteria
Language	Studies published in English	Studies published in languages other than English
Publication period	Studies published between 2017 and 2026	Studies published before 2017
Publication type	Peer-reviewed journal articles	Conference proceedings, conference papers, books, book chapters, reviews, and other non-journal publications
Publication stage	Final published articles	Articles classified as in press or not yet in final publication form
Subject area	Studies indexed within Social Sciences and/or Arts and Humanities	Studies indexed exclusively outside Social Sciences and Arts and Humanities
Access status	Full-text articles available through open access	Articles not available through open access

Eligibility Assessment and Inclusion

The 69 unique records underwent title and abstract screening, followed by full-text eligibility assessment. Records were excluded if they fell outside the scope of the review, did not address text complexity or related linguistic characteristics of children's storybooks, or did not provide

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

sufficient full-text information for assessment. 54 records (78.3%) were excluded. 15 records (21.7%) met the eligibility criteria and proceeded to the quality appraisal.

PRISMA framework (Page et al., 2021) was used to ensure a transparent and systematic study selection process. Starting with 443 records found in Scopus and WoS, 348 records were excluded during the screening process, leaving behind 95 records. 26 duplicates were found and removed, and the remaining 69 unique records underwent eligibility assessment. A further 54 records were excluded for not adhering to the eligibility requirements, which left 15 records that proceeded to quality appraisal and subsequent synthesis. Figure 1 presents the complete study selection process.

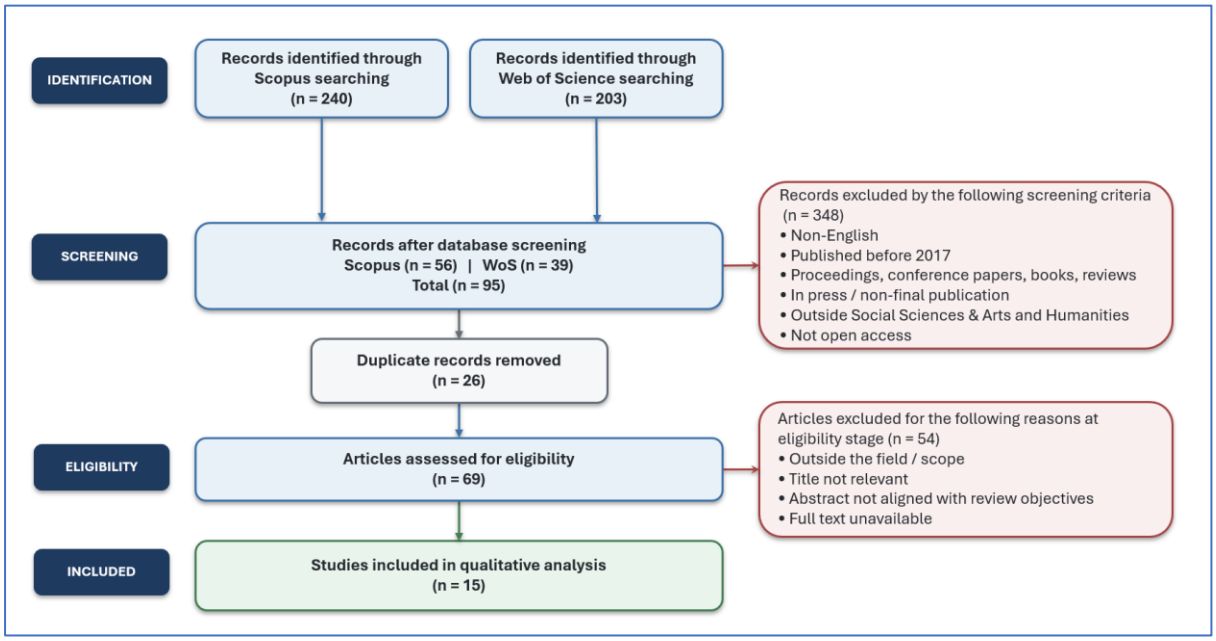


Figure 1. PRISMA Flow Diagram of the Study Selection Process

Thematic Analysis

The thematic analysis involved an in-depth examination of the selected studies. These 15 studies were examined based on a predefined framework aligned with the review objectives and research questions. The extracted information included research objectives, methodological designs, analytical procedures, measures of text complexity and other principal findings. The information was then organised into three corresponding themes: i) Computational Measures of Text Complexity and Readability, ii) Linguistic Dimensions of Text Complexity, and iii) Reader-Text Interaction and Text Complexity. Each study was first placed under the theme that best reflected its main contribution. Where a study also provided evidence relevant to another theme, it was included there as well. The four senior authors reviewed the grouping of the studies and refined the organisation of the findings so that the themes remained consistent with the research questions and the evidence reported in the selected studies. Differences in interpretation were discussed until agreement was reached. The final synthesis followed the three research questions: how text complexity is measured, which linguistic features contribute to it, and how children experience these demands through their

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

interaction with the text. This structure was used to organise and interpret the findings across the studies. Table 3 shows the thematic classification of the included studies.

Table 3. Thematic Classification of Studies on Text Complexity in Children's Storybooks

Theme	Focus Area	Main Research Direction
1. Computational Measures of Text Complexity and Readability	Readability indices, computational linguistic measures, lexical and syntactic indices, corpus-based analysis, automated text analysis, and machine-learning approaches	Examines how quantitative and computational measures are used to assess, differentiate, and predict the complexity and readability of children's texts across age or difficulty levels.
2. Linguistic Dimensions of Text Complexity	Lexical, morphological, syntactic, grammatical, and discourse-functional features, including linguistic adaptation, simplification, and translation	Investigates which linguistic characteristics contribute to text complexity and how changes in lexical, syntactic, grammatical, and discourse features influence the readability and accessibility of children's texts.
3. Reader-Text Interaction and Text Complexity	Reader characteristics, language experience, shared reading, comprehension, adult or technological mediation, illustrations, text–image relationships, multimodal processing, and mode of presentation	Examines text complexity as an interaction between textual features and readers, including how developmental and linguistic resources, reading conditions, social interaction, and multimodal features influence the experience and comprehension of children's texts.

Quality Assessment Analysis

Following the recommendations of Kitchenham and Charters (2007), the methodological quality of the primary studies was assessed after the study selection process to determine the strength and credibility of the evidence included in the review. For this purpose, the quality assessment framework proposed by Abouzahra et al. (2020) was adopted. The framework consists of six quality assessment (QA) criteria designed to provide a systematic and consistent evaluation of the selected studies. Each criterion was rated using a three-level scoring system. A score of 1 (Yes) was assigned when the criterion was fully satisfied, 0.5 (Partly) when the criterion was addressed but demonstrated some limitations or insufficient detail, and 0 (No) when the criterion was not satisfied, or the required information was not reported. The resulting scores provided a structured basis for comparing the methodological quality of the included studies and assessing the overall strength of the evidence contributing to the review.

QA1. Is the purpose of the study clearly stated?
QA2. Is the interest and the usefulness of the work clearly presented?
QA3. Is the study methodology clearly established?
QA4. Are the concepts of the approach clearly defined?
QA5. Is the work compared and measured with other similar work?
QA6. Are the limitations of the work clearly mentioned?

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

The quality assessment (QA) used a predefined set of criteria to evaluate the methodological quality and reporting adequacy of the studies included in the review. Four senior authors served as experts, independently evaluated each study against the six QA criteria, and assigned one of three ratings: Yes (Y = 1), Partly (P = 0.5), or No (N = 0). The experts assessed each study independently to reduce reliance on a single evaluator's judgement and strengthen the consistency of the appraisal process. Following the independent assessments, differences in ratings were discussed until consensus was reached on the final score for each criterion and study. A minimum threshold was established for inclusion in the subsequent synthesis, whereby only studies achieving an aggregated quality score above 3.0 were retained. Studies scoring at or below this threshold were excluded because they did not meet the specified quality criteria. This procedure provided a transparent, systematic mechanism for distinguishing studies with adequate methodological and reporting quality from those with substantial limitations, strengthening the evidential basis of the final synthesis.

The quality assessment results demonstrate a consistently strong methodological and reporting profile across the 15 selected studies. Eight (53.3%) achieved the maximum score of 6.0 (100%), indicating full compliance with all six quality assessment criteria. Six studies (40.0%) scored 5.5 (91.67%), and one study (6.7%) scored 4.5 (75%). Overall, quality scores ranged from 75% to 100%, with a mean of approximately 5.70 out of 6.0 (95.0%), indicating strong adherence to the predefined QA criteria across the evidence base. Notably, none of the studies fell below the established inclusion threshold of 3.0. The lower scores were primarily attributable to criteria that were only partially satisfied or not satisfied, particularly QA5 and QA6, rather than weaknesses across all assessment dimensions. Collectively, these findings indicate that the studies retained for synthesis had generally strong methodological and reporting quality, providing a credible evidence base for addressing the objectives of the systematic literature review. Table 4 displays the quality appraisal of the 15 selected studies.

Table 4: Quality Appraisal of the 15 Selected Studies

No.	Authors	Article Title	QA 1	QA 2	QA 3	QA 4	QA 5	QA 6	Total /6	%
1	Alruthaya et al. (2025)	Moving towards a more comprehensive understanding of multimodal text complexity	Y	Y	Y	Y	Y	P	5.5	91.67
2	Green and Keogh, (2024)	Vocabulary exposure to children is enhanced by using both informational and narrative picture books	Y	Y	Y	Y	Y	P	5.5	91.67

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

No.	Authors	Article Title	QA 1	QA 2	QA 3	QA 4	QA 5	QA 6	Total /6	%
		for read-alouds								
3	Diehm et al. (2020)	Young children's narrative retell in response to static and animated stories	Y	Y	Y	Y	Y	Y	6.0	100.00
4	Gök et al. (2021)	Feridun Oral's children's books' peritextual features and their text readability levels	Y	Y	Y	Y	P	N	4.5	75.00
5	Green et al. (2024)	The Children's Picture Books Lexicon (CPB-Lex)	Y	Y	Y	Y	Y	Y	6.0	100.00
6	Yang et al. (2025)	Chinese mothers use idioms in shared book reading	Y	Y	Y	Y	Y	Y	6.0	100.00
7	Han (2025)	Normalization or creation?	Y	Y	Y	Y	Y	Y	6.0	100.00
8	Liu (2024)	Readability and adaptation of children's literature	Y	Y	Y	Y	Y	P	5.5	91.67
9	Montag (2019)	Differences in sentence complexity in the text of children's picture books and child-directed speech	Y	Y	Y	Y	Y	P	5.5	91.67

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

No.	Authors	Article Title	QA 1	QA 2	QA 3	QA 4	QA 5	QA 6	Total /6	%
10	Nojonen et al. (2026)	Characterizing Finnish children's books by target age	Y	Y	Y	Y	Y	Y	6.0	100.00
11	Parlindungan et al. (2024)	Evaluating Indonesian Hero Stories Featured in Children's Literature	Y	Y	Y	Y	Y	Y	6.0	100.00
12	Ślęzak-Świat (2026)	Pre-digitally decoded Alice's Adventures	Y	Y	Y	Y	P	Y	5.5	91.67
13	Wang (2017)	Interpreting the Readability in Simplified Editions of Children's Literature	Y	Y	Y	Y	Y	P	5.5	91.67
14	Xu et al. (2021)	Same benefits, different communication patterns	Y	Y	Y	Y	Y	Y	6.0	100.00
15	Zhao et al. (2022)	Linguistic Indicators for Text Complexity in Picture Books	Y	Y	Y	Y	Y	Y	6.0	100.00

Results and Discussion

Computational Measurement of Text Complexity and Readability

Traditional readability formulas continue to be used in studies of children's storybooks, although more recent research has increasingly incorporated computational measures that draw on multiple indicators of text complexity. Readability formulas are particularly useful for making numerical comparisons between texts or between versions of the same text. For example, Gök et al. (2021) examined children's storybooks using word length, sentence length and readability formulas. Their analysis showed considerable variation among books by the same author. The variation suggests that storybooks from similar literary contexts cannot necessarily be assumed to have the same level of readability. Wang (2017) similarly reported Flesch Reading Ease scores of 80.4 for the original *Alice's Adventures in Wonderland*, 84.6 for

Green's adaptation, and 94.7 for Swan's adaptation. These results demonstrate the usefulness of readability formulas for quantifying differences between texts. However, such formulas provide limited information about the particular textual characteristics responsible for those differences. This limitation has encouraged the use of computational measures that can process a wider range of linguistic information. Zhao et al. (2022), for instance, automatically analysed eleven linguistic indicators in 29 picture books and found significant differences across kindergarten 1–3 levels in seven indicators: frequency of tokens, frequency of types, vocabulary diversity, mean length of utterance in morphemes, number of verbs per utterance, number of DSS-eligible utterances, and total number of non-zero morphemes.

A second development is the use of computational tools to generate multiple measures systematically rather than relying on a small number of manually calculated indices. Zhao et al. (2022) used CLAN to automatically analyse the texts of children's picture books. The computational procedure examined vocabulary and syntactic indicators and compared the resulting measures with experts' overall levelling of the books. This approach is important because computational analysis provides objective and replicable information that can assist the comparison and classification of texts. However, Zhao et al. (2022) caution that automated analysis should not function as the sole method for grading picture books. Elements such as illustrations, layout, rhyme, prior knowledge, and reading experience fall outside many text-based measures. Green et al. (2024) demonstrate a different form of computational measurement through the development of the Children's Picture Books Lexicon (CPB-Lex). The database was constructed from approximately 1.12 million tokens across 2,146 picture-book readings and contains approximately 25,585-word types. It provides frequency norms and additional lexical information that can support detailed computational analysis of children's print exposure. Rather than reducing measurement to a single readability value, this type of resource allows multiple lexical statistics to be derived and compared across a large collection of texts. Computational measurement therefore expands both the number of texts that can be analysed and the range of indicators that can be incorporated into assessments of complexity.

Green and Keogh (2024) further demonstrate the application of data-driven methods to the analysis of vocabulary exposure in children's picture books. Using comparative modelling, the study examined informational and narrative picture books used for read-alouds to determine how the two text types contribute to children's vocabulary exposure. The findings showed that the two genres provide complementary lexical opportunities, suggesting that computational analysis can reveal differences in vocabulary characteristics that may not be captured through conventional readability measures alone. This study extends computational approaches to text complexity by demonstrating how data-science methods can be used to compare lexical properties across different types of children's books.

Ślęzak-Świat (2026) examined *Alice's Adventures in Wonderland* and *Through the Looking-Glass using several measures. The study used* mean length of utterance, lexical profiling, and the Gunning Fog Index. These measures captured different aspects of textual complexity and did not rely on a single readability score. Together, they allow examination of differences in vocabulary, sentence structure and readability. This therefore shows that a readability score may not be enough to explain the complexity of children's literary texts.

The movement towards multidimensional computational measurement is most clearly demonstrated by Nojonen et al. (2026), who analysed 300 Finnish children's books using 148 linguistic features. The study incorporated traditional readability indices alongside larger sets

of automatically extracted features and applied computational procedures to determine which characteristics differentiated books intended for different target ages. The traditional measures included Flesch–Kincaid Grade Level, a modified SMOG index for Finnish, the Coleman–Liau Index, and the Automatic Readability Index. However, these measures formed only one component of a much larger analytical framework. Key Feature Analysis was used to identify characteristics associated with text classes, while a Support Vector Machine provided a predictive classification of books according to target-age groups. The machine-learning procedure used a one-versus-one classification strategy, with normalisation and author-level restrictions introduced to reduce the possibility that the model would learn individual authorial styles rather than characteristics associated with the text categories. This approach does more than give a numerical score for text difficulty. It examines whether several linguistic features together can distinguish books for different age groups. Nojonen et al. (2026) also used FinBERT as a deep-learning baseline. However, knowing that a model can classify books is not enough. Linguistic features remain important because they help explain what makes the texts different in terms of complexity.

Overall, the studies used different computational approaches to measure text complexity. Some relied on traditional readability formulas, while others used automated linguistic measures, corpus data, or machine-learning methods. Gök et al. (2021) and Wang (2017), for example, used readability measures to compare books or different versions of a text. Other studies examined a wider range of information. Zhao et al. (2022) combined linguistic measures with expert judgements of book difficulty, and Green et al. (2024) developed a large lexical resource based on children's picture books. Green and Keogh (2024) examined vocabulary exposure, while Ślęzak-Świat (2026) used several quantitative measures to examine literary texts. Nojonen et al. (2026) went further by analysing a large set of linguistic features and using machine learning to classify books according to target age. Taken together, these studies suggest that text complexity cannot be fully measured with a single method. Readability formulas are useful for giving an overall estimate, but automated and corpus-based methods can provide information on a wider range of textual features. Machine-learning methods add another possibility by examining patterns across many features at the same time. In this review, these studies are grouped together because they focus mainly on how text complexity is measured. The linguistic features that contribute to complexity are discussed separately under *Linguistic Dimensions of Text Complexity*.

Linguistic Dimensions of Text Complexity

Children's storybooks differ in more than their word and sentence length. Differences can also be found in vocabulary, morphology, syntax, grammar, and discourse. Zhao et al. (2022), for example, examined eleven linguistic features in 29 picture books across three kindergarten levels. Seven features differed significantly across the levels: token frequency, type frequency, vocabulary diversity, mean length of utterance in morphemes, verbs per utterance, DSS-eligible utterances, and the total number of non-zero morphemes. This shows that differences between book levels involve several aspects of language rather than one feature alone. Gök et al. (2021) reported considerable differences across the 13 storybooks they examined. One book, for example, had sentences averaging 15.65 words and a readability score of 48.41. In *The Hedgehog and the Chestnut*, the average was only 6.15 words per sentence, with a readability score of 86.32. Clearly, the two texts differ in readability. However, these scores cannot explain what aspects of the language account for the difference. This is where linguistic analysis becomes useful. Liu (2024) examined one of these features, interpersonal grammatical

metaphor, with particular attention to mood and modality. Readability scores may show that texts differ in difficulty, but linguistic analysis can help explain the source of those differences. Vocabulary also contributes to differences in text complexity. Zhao et al. (2022) found that picture books for kindergarten 1 to kindergarten 3 varied in word types, word tokens, and lexical diversity. The differences across levels were not only related to the number of words used, but also to the range of vocabulary. A similar pattern was reported by Wang (2017), where simplified versions of *Alice's Adventures in Wonderland* contained shorter words and less lexical variation than the original text. Nojonen et al. (2026) also examined lexical variation and morpho-lexical richness in 300 Finnish children's books written for different age groups. Across these studies, vocabulary complexity involves more than the presence of difficult words. The amount of vocabulary, its variation, and the forms of words used in a story can all make a difference. How often words occur, how much the vocabulary varies, and the morphological structure of words can all contribute to complexity. A story may therefore use mostly familiar words but still place greater demands on readers when the vocabulary is highly varied or the word forms are more complex.

Sentence structure and grammar also contribute to the complexity of children's storybooks. Montag (2019) found that passive constructions and several types of relative clauses occurred more often in picture books than in child-directed speech. This suggests that children may encounter grammatical structures in books that they hear less often in everyday conversation. Differences were also found across picture-book levels. Zhao et al. (2022) reported variation in verbs per utterance, Developmental Sentence Scoring-eligible utterances, mean length of utterance in morphemes, and total non-zero morphemes. These differences cannot be explained by sentence length alone because they also involve the structure and grammatical content of sentences. Wang (2017) found several grammatical changes in simplified versions of *Alice's Adventures in Wonderland*. These included fewer clause complexes, shorter clauses, changes in clause order, and simpler theme structures. Marked themes, for example, decreased from 13.0% in Carroll's original to 2.1% in Swan's adaptation, while unmarked themes reached 89.0%. Nojonen et al. (2026) approached sentence structure differently by separating clausal and phrasal complexity. Such differences are not always captured by sentence length. Two sentences of similar length may differ considerably in how their clauses and phrases are structured. Examining these structures can therefore reveal aspects of grammatical difficulty that sentence length alone may miss.

How meaning is organised within a story can also add to its complexity. Liu (2024) examined mood and modality in adapted children's literature through interpersonal grammatical metaphor. Three strategies were identified: unpacking, demetaphorization, and maintenance. The way meaning is expressed therefore matters, especially when it is indirect or metaphorical. Wang (2017) found another type of change in simplified versions of *Alice's Adventures in Wonderland*. Changes were also made to theme choice and clause organisation, affecting how information was introduced and developed. This shows that simplification involves more than replacing difficult words or shortening sentences. It can also shape how ideas are organised and linked across the text.

A story does not always remain linguistically the same when it is adapted or translated. Liu (2024), for instance, found changes in grammar and discourse in adaptations of children's literature. Wang (2017) reported changes in vocabulary, clauses, and theme organisation in simplified versions of *Alice's Adventures in Wonderland*. Translation may change the language of a story in other ways. In Han's (2025) corpus-based study, Chinese translations of English

children's texts showed patterns of normalisation towards the conventions of Chinese. This means that complexity may change as a story moves from one version or language to another. Some linguistic features may be simplified, reorganised, or adjusted rather than simply retained from the original.

Overall, no single linguistic feature explains the complexity of a children's storybook. Vocabulary, morphology, grammar, sentence structure, and the way information is organised may all play a part. Their contribution can also vary from one text to another. For this reason, examining only one feature may give an incomplete picture of the language children encounter in a story. In this theme, the concern is with the language itself: what features make one storybook linguistically more demanding than another. This differs from measuring complexity through computational methods or examining how children respond to a text. Considering these linguistic features together may therefore be more useful when storybooks are compared, selected, or adapted for young readers.

Reader-Text Interaction and Text Complexity

The reviewed studies indicate that the experience of text complexity is shaped not only by characteristics within a storybook but also by the relationship between textual demands and the resources and conditions through which children engage with the text. Children may encounter the same story differently depending on the linguistic input available during reading, the support provided by adults or technological partners, and the way verbal and visual information is presented. Yang et al. (2025), for example, examined shared book reading among 33 Chinese children aged 3–5 years in Norway using a wordless picture book. Maternal idiom uses during the activity predicted growth in children's receptive Chinese vocabulary over one year, although the same relationship was not found for expressive vocabulary. Mothers' language use varied considerably: 22 mothers collectively used 27 different idioms across 60 occurrences, whereas 11 mothers did not use idioms. Maternal idiom use was also associated with the amount and lexical diversity of maternal talk. These findings show that the linguistic experience associated with a storybook extends beyond the language in the book itself. The interaction surrounding a story can introduce additional vocabulary and expressions, thereby influencing the linguistic resources available to children as they construct meaning.

Social and conversational mediation represents a particularly important dimension of reader-text interaction. Yang et al. (2025) show that adults can expand the linguistic environment around a story through their language choices, while Xu et al. (2021) show that the nature of the reading partner can influence how children participate in story-related conversation. In their study of 90 children aged 3–6 years, Xu et al. (2021) compared guided storybook conversation with a smart-speaker conversational agent to interaction with a human partner. The conversational agent supported story comprehension at a level comparable to that of the human partner, but differences emerged in children's verbal responses. Children interacting with a human partner produced longer, more lexically diverse, and more topically relevant responses, whereas interaction with the conversational agent was associated with greater intelligibility. These differences suggest that similar comprehension outcomes do not necessarily indicate identical reading experiences. The form of mediation may influence the amount and characteristics of language children produce while responding to a story. Reader-text interaction should therefore be understood as involving not only the relationship between the child and the written text but also the social or technological support through which engagement with that text takes place.

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

The mode of story presentation further shows that reading conditions shape the experience of textual demands. Diehm et al. (2020) compared young children's narrative retells following static and animated stories and found differences in several measures of their responses. Children produced an average of 46.77 words following the animated condition compared with 34.01 following the static-book condition, while the number of different words increased from 19.30 to 24.52. Mean length of utterance was also greater following animation, and fewer examiner prompts were required. After adjustment for multiple comparisons, significant differences remained for total words, different words, mean length of utterance, action verbs, and examiner prompts. The way a story is presented can also shape how children respond to it. Even when the wording stays the same, children may engage with the story differently depending on the format or mode of presentation. This means that the difficulty they experience cannot always be explained by the language of the text alone.

This interaction becomes particularly important in picture books, where meaning is constructed through relationships between written and visual information. Alruthaya et al. (2025) found that children encountered unexpected difficulties when information presented through illustrations was inconsistent with or complicated meanings conveyed through print. Importantly, even apparently simple and high-frequency words could become difficult when visual information introduced competing or additional meanings. Eye-movement and miscue evidence also revealed differences in how readers coordinated their attention between verbal and visual information. Some readers initially focused on print but later showed weaker understanding of visual-verbal relationships in their retelling, whereas others returned to illustrations during reading in ways that supported meaning construction. These findings suggest that complexity may emerge from the need to integrate information across modes rather than from either written language or illustrations alone. For picture books in particular, reader-text interaction therefore includes the reader's ability to coordinate verbal and visual information, a dimension that conventional readability measures based on words and sentences are unlikely to capture adequately.

Reader-text suitability is also reflected in the judgments made by adults responsible for selecting materials for children. Parlindungan et al. (2024) examined 25 elementary-school teachers' evaluations of three Indonesian hero storybooks intended for Grades 4–6. Their evaluations considered text complexity, plot, illustrations, moral values, and the books' suitability for educational purposes. Perceptions varied across the three books. One book was generally considered interesting and accessible because of its clear plot, whereas the other books received more divided evaluations because of characteristics such as lengthy descriptions, complexity, and narrative twists. These findings indicate that judgments about the appropriateness of storybooks are not based exclusively on measurable linguistic characteristics. Teachers interpret textual and visual characteristics in relation to the perceived abilities, interests, and educational needs of intended readers. Their role is another link between the text and the reader, especially since adults often decide which storybooks children are exposed to before reading begins.

Overall, the difficulty of a storybook is not determined by the text alone. Children's experiences can also be shaped by how they interact with the story and the setting in which reading takes place. Yang et al. (2025) and Xu et al. (2021), for instance, show how social and technological support can influence storybook activities. Diehm et al. (2020) also reported differences in children's narrative responses when the same story was presented in different ways. Alruthaya et al. (2025), meanwhile, found that difficulty could arise when children had to connect

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228
information from words and images. Judgements about whether a book is suitable for children may also depend on the intended readers, as reported by Parlindungan et al. (2024). These findings add something that text measures alone cannot capture. A book may have particular linguistic features, but the demands experienced by children can also change with the way the story is presented, supported, and understood. Text complexity, therefore, needs to be considered in relation to both the storybook and the conditions under which children engage with it.

Conclusion

This review examined studies on text complexity in children's storybooks published from 2017 to 2026. The studies were identified through searches in Scopus and Web of Science and selected following the PRISMA framework and the criteria set for the review. Only peer-reviewed, final-stage, open-access journal articles in English from the Social Sciences and Arts and Humanities were considered. The analysis addressed three questions. The first looked at how text complexity and readability have been measured. The second examined the linguistic features associated with complexity, while the third considered what happens when children interact with storybooks under different reading conditions. These questions allowed text complexity to be considered from three related angles: how it is measured, what features of the text contribute to it, and how those demands are encountered by readers.

The first finding relates to how text complexity and readability are measured. Traditional readability formulas are still commonly used because they provide numerical scores based mainly on word and sentence characteristics. However, the studies reviewed here did not rely on these scores alone. Several used automated linguistic analysis, corpus data, or machine-learning methods to examine a wider range of text features. Large lexical databases also made it possible to study vocabulary in children's texts in greater detail. Some recent studies combined many linguistic features to distinguish texts written for different ages or levels of difficulty. This does not mean that traditional readability formulas are no longer useful. They provide a general indication of difficulty, while newer computational methods can examine patterns that a single score may not capture. The studies therefore point to the value of using different measures together when examining the complexity of children's storybooks.

The second finding concerns the language used in children's storybooks. Complexity may come from vocabulary, morphology, sentence structure, grammar, or the way meaning is organised in a text. For vocabulary, difficulty is not limited to unfamiliar words. Word frequency, vocabulary diversity, distribution, and morphological features also matter.

Sentence length by itself does not explain why a text is difficult. A long sentence may be challenging, but difficulty can also come from the way clauses are arranged, the use of subordination, passive forms, phrases, and relative clauses. Problems may also arise from how ideas are connected from one sentence to the next. Research on simplified and adapted storybooks shows that easier versions often involve several kinds of changes rather than just shorter sentences or simpler words. This suggests that storybook complexity is usually the result of several features working together.

The third finding looks at what children bring to the reading experience. The same story may not be equally difficult for every child. Vocabulary knowledge, language ability, and previous exposure to language can affect how much they understand. What happens while reading is

also important. In shared reading, for example, an adult may explain a word, ask a question, or provide information that is not in the book. Children's responses can also differ depending on who reads with them. The format of the story adds another factor. Pictures, animation, and spoken language may support understanding, but they can also change what children need to process. A child may understand the words in a sentence but still find it difficult to connect them with an illustration. Difficulty, then, comes not only from the text itself but also from the child's knowledge and the way the story is experienced.

This review draws together three areas that are often examined separately: how text complexity is measured, which language features make a text difficult, and how children experience that difficulty when reading or listening to a story. Bringing these areas together shows the limits of relying on readability scores alone. Such scores can indicate that one book is more difficult than another, but they do not always explain why. They also cannot show whether every child will experience the same level of difficulty. Linguistic analysis can help identify the features that make a text demanding, while studies of children's responses show how difficulty may vary from one reader or reading situation to another. Looking at these areas together gives a clearer picture of text complexity in children's storybooks.

These findings may be useful when storybooks are chosen, written, or adapted for children. Readability scores and age labels can give a general idea of difficulty, but they do not show everything. A closer look at the vocabulary, grammar, and organisation of the text may reveal problems that the score does not pick up. The children who will use the book also need to be considered, particularly their language ability and previous reading experience. For authors and publishers, these factors may be useful when deciding what to change in a story intended for a different age or proficiency level. Importantly, visual presentation and reading conditions should be considered where storybooks depend substantially on interactions among print, illustration, spoken language, and other modes. Evaluation of children's texts can therefore benefit from combining objective textual evidence with consideration of how intended readers are likely to encounter and process that information.

Several limitations should be acknowledged. The review was restricted to Scopus and Web of Science, English-language publications, and the period from 2017 to 2026, which may exclude relevant research available through other databases, languages, or earlier periods. The inclusion criteria relating to publication type, subject area, publication stage, and open-access availability may have further limited the evidence base. Future research should examine larger and more linguistically and culturally diverse collections of children's storybooks. Further work is also needed to link computational estimates to children's actual comprehension and reading behaviour and to determine which linguistic measures best represent meaningful differences in complexity. Research examining how textual characteristics interact with reader proficiency, social mediation, illustrations, and digital presentation would further strengthen understanding of reader-text relationships.

Overall, this review shows that text complexity in children's storybooks is best understood through the relationship among measurement, linguistic characteristics, and reader-text interaction. Computational approaches provide increasingly sophisticated ways to assess complexity, linguistic analysis explains the textual sources of that complexity, and reader-oriented evidence shows that the demands of a storybook vary with the resources and reading conditions surrounding the child. This integrated perspective provides a stronger basis for

future empirical research and for evidence-based selection, evaluation, adaptation, and development of storybooks appropriate for young readers.

Acknowledgements:	The authors express sincere gratitude to Universiti Teknologi MARA for providing the necessary resources and support throughout this research. Special appreciation is extended to colleagues and peers who contributed valuable insights and constructive feedback, which greatly enhanced the quality of this paper.
Funding Statement:	This work was supported by funding from the Ministry of Higher Education Malaysia through the Akademi Pengajian Bahasa, Universiti Teknologi MARA, Strategic Matching Grant UiTM.800-3/1/21. The funding body had no role in the study design, data collection, analysis, interpretation of results, or the decision to publish this manuscript.
Conflict of Interest Statement:	The authors declare no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to this journal.
Ethics Statement:	This study did not involve any human participants, animals, or sensitive data requiring ethical approval. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.
Author Contribution Statement:	All authors contributed significantly to this manuscript. Anealka Aziz led the conceptualisation and methodology of the study and was responsible for data analysis, interpretation of the results, and overall supervision. Suzana and Muna Liyana collected data and assisted with analysis and interpretation. Ng Yu Jin, Najiha, and Mallika contributed to the literature review and critical revision of the manuscript. Anealka Aziz, Muna Liyana, Ng Yu Jin, and Mallika conducted the quality assessment of the articles included in the systematic literature review (SLR). All authors read and approved the final version of the manuscript prior to submission.

International Journal of Education, Psychology and Counseling
IJEPC
EISSN : 0128-164X
Volume 11 Issue 65 (December 2026) PP. 207-228

References

Abouzahra, A., Sabraoui, A., & Afdel, K. (2020). Model composition in Model Driven Engineering: A systematic literature review. *Information and Software Technology*, *125*(May), 106316. https://doi.org/10.1016/j.infsof.2020.106316

Alruthaya, S., Mantei, J., White, S. L. J., & Kervin, L. (2025). Moving towards a more comprehensive understanding of multimodal text complexity. *Australian Journal of Language and Literacy*, *48*(2), 119–139. https://doi.org/10.1007/s44020-025-00079-9

Andreeva, M., Solnyshkina, M., & Saadna, S. (2025). Complexity Mechanism in Leveled Texts. *RUDN Journal of Language Studies, Semiotics and Semantics*. https://doi.org/10.22363/2313-2299-2025-16-3-821-844

Balaban, H. Ö., & Hohenberger, A. (2020). The development of narrative skills in Turkish-speaking children: A complexity approach. *PLoS ONE*, *15*(5). https://doi.org/10.1371/journal.pone.0232579

Battistelli, D., Etienne, A., Rahman, R., Teissèdre, C., & Lecorvé, G. (2022). A Processing Chain to Explain the Complexity of Texts for Children From a Linguistic and Psycho-linguistic Point of View. In E. Y., J. T., P. T., & B. M.Z. (Eds.), *Traitement Automatique des Langues Naturelles, TALN 2022 - Actes de la 29e Conference sur le Traitement Automatique des Langues Naturelles: Conference Principale* (Vol. 1, pp. 236–246). Association pour le traitement automatique des langues. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85175851620&partnerID=40&md5=7be0a875ff007e2368ba6ea7636d4239

Diehm, E. A., Wood, C., Puhlman, J., & Callendar, M. (2020). Young children's narrative retell in response to static and animated stories. *International Journal of Language and Communication Disorders*, *55*(3), 359–372. https://doi.org/10.1111/1460-6984.12523

Fitzgerald, J., Koons, H., Bowen, K., Sanford-Moore, E. E., Elmore, J., Hiebert, E. H., & Jackson Stenner, A. (2015). Important text characteristics for early-grades text complexity. *Journal of Educational Psychology*, *107*(1), 4–29. https://doi.org/10.1037/a0037289

Gibbs, A., & Reed, D. (2024). Language Development Opportunities in Curricular and Trade Books: A Content Analysis. *European Journal of Education*. https://doi.org/10.1111/ejed.12807

Gök, B., Temizyurek, F., Bas, O., & Sirem, O. (2021). Feridun Oral's children's books' peritextual features and their text readability levels. *TURKISH JOURNAL OF EDUCATION*, *10*(2), 125–138. https://doi.org/10.19128/turje.808967

Green, C., & Keogh, K. (2024). Vocabulary exposure to children is enhanced by using both informational and narrative picture books for read-alouds: A comparative modelling study using data science methods. *Journal of Research in Reading*, *47*(4), 497–516. https://doi.org/10.1111/1467-9817.12462

Green, C., Keogh, K., Sun, H., & O'Brien, B. (2024). The Children's Picture Books Lexicon (CPB-Lex): A large-scale lexical database from children's picture books. *Behavior Research Methods*, *56*(5), 4504–4521. https://doi.org/10.3758/s13428-023-02198-y

Günaydın, Y. (2023). PIKTES İÇİN HAZIRLANAN HİKÂYE SERİLERİNİN OKUNABİLİRLİK DÜZEYLERİ. *Sakarya University Journal of Education*. https://doi.org/10.19126/suje.1296885

Han, Y. (2025). Normalization or creation? A corpus-based study of normalization in the Chinese translation of English children's literature. *Humanities and Social Sciences Communications*, *12*(1). https://doi.org/10.1057/s41599-025-05379-6

Hartati, E., Nababan, M., Santosa, R., & Djatmika. (2025). The Readability of the Indonesian Translation of English Storybook's Lexis: A Perspective from the Adults Vs. Children Readers. *Forum for Linguistic Studies*. https://doi.org/10.30564/fls.v7i11.11413

Hiebert, E. (2025). Flattening the Developmental Staircase: Lexical Complexity Progression in Elementary Reading Texts Across Six Decades. *Education Sciences*. https://doi.org/10.3390/educsci15111546

Hsiao, Y., Dawson, N. J., Banerji, N., & Nation, K. (2024). A corpus-based developmental investigation of linguistic complexity in children's writing. *Applied Corpus Linguistics*, *4*(1). https://doi.org/10.1016/j.acorp.2024.100084

Imperial, J. M., & Ong, E. (2021). Application of Lexical Features Towards Improvement of Filipino Readability Identification of Children's Literature. *ArXiv*, *abs/2101.1*. https://doi.org/10.48550/arxiv.2101.10537

Kawar, K., Saiegh-Haddad, E., & Armon-Lotem, S. (2023). Text complexity and variety factors in narrative retelling and narrative comprehension among Arabic-speaking preschool children. *First Language*, *43*, 355–379. https://doi.org/10.1177/01427237221149800

Kitchenham, B. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical Report, Ver. 2.3 EBSE Technical Report. EBSE*.

Kupriyanov, R., Bukach, O., & Aleksandrova, O. (2023). Cognitive complexity measures for educational texts: Empirical validation of linguistic parameters. *Russian Journal of Linguistics*. https://doi.org/10.22363/2687-0088-35817

Liu, M., Li, Y., Su, Y., & Li, H. (2023). Text Complexity of Chinese Elementary School Textbooks: Analysis of Text Linguistic Features Using Machine Learning Algorithms. *Scientific Studies of Reading*, *28*, 235–255. https://doi.org/10.1080/10888438.2023.2244620

Liu, Y. (2024). Readability and adaptation of children's literature: An interpersonal metaphor perspective. *Journal of World Languages*, *10*(3), 601–622. https://doi.org/10.1515/jwl-2022-0039

Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis: Methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, *13*(3), 179–187. https://doi.org/10.1097/XEB.0000000000000062

Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, *39*(5), 527–546. https://doi.org/10.1177/0142723719849996

Muhinyi, A., Stewart, A., & Rowland, C. (2025). Encouraging use of complex language in preschoolers: a classroom-based storybook intervention study. *Language Learning and Development*, *21*, 399–417. https://doi.org/10.1080/15475441.2024.2443447

Nojonen, T., Kanerva, J., Bertram, R., Laakso, M. J., Ginter, F., & Laippala, V. (2026). Characterizing Finnish children's books by target age - Complementing key feature analysis with machine learning. *APPLIED CORPUS LINGUISTICS*, *6*(3). https://doi.org/10.1016/j.acorp.2026.100240

Page, M. J., McKenzie, J. E., Bossuyt, P., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The prisma 2020 statement: An updated guideline for reporting systematic reviews. *Medicina Fluminensis*, *57*(4), 444–465. https://doi.org/10.21860/medflum2021_264903

Parlindungan, F., Rifai, I., Nuthihar, R., & Dewayani, S. (2024). EVALUATING INDONESIAN HERO STORIES FEATURED IN CHILDREN'S LITERATURE. *LLT Journal: Journal on Language and Language Teaching*, *27*(2), 931–947. https://doi.org/10.24071/llt.v27i2.8787

Petrie, A., Mayr, R., Zhao, F., & Montanari, S. (2021). Parent-child interaction during storybook reading: wordless narrative books versus books with text. *Journal of Child Language*, *50*, 104–131. https://doi.org/10.1017/s0305000921000763

Schwarz, A. L., Gonzales, M., Melamed, T., & Abdi, H. (2024). How do speech language pathologists use the Schwarz et al. (2015) rubric for read-aloud storybook selection? *Speech, Language and Hearing*, *28*. https://doi.org/10.1080/2050571x.2024.2404315

Ślęzak-Świat, A. (2026). Pre-digitally decoded alice's adventures: an early record of readability computations. *Digital Scholarship in the Humanities*, *41*(1), 412–419. https://doi.org/10.1093/llc/fqaf146

Tortorelli, L. S. (2020). Beyond first grade: examining word, sentence, and discourse text factors associated with oral reading rate in informational text in second grade. *Reading and Writing*, *33*(1), 143–170. https://doi.org/10.1007/s11145-019-09956-5

Wang, T. (2017). Interpreting the readability in simplified editions of children's literature from a thematic perspective: A case study on Alice's Adventures in Wonderland. *Indonesian Journal of Applied Linguistics*, *7*(1), 201–205. https://doi.org/10.17509/ijal.v7i1.6876

Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing Children's reading with a conversational agent vs. a human partner. *Computers and Education*, *161*. https://doi.org/10.1016/j.compedu.2020.104059

Yang, J. Y., Grover, V., & Lawrence, J. F. (2025). Chinese mothers use idioms in shared book reading: A predictor for children's Chinese vocabulary growth? *JOURNAL OF CHILD LANGUAGE*, *52*(4), 918–944. https://doi.org/10.1017/S0305000924000266

Zhao, J., Zhu, M., De Ruiter, L., & Chen, S. (2022). *Linguistic Indicators for Text Complexity in Picture Books for Young Chinese Children Learning English as a Foreign Language*. *7*. https://doi.org/10.3389/feduc.2022.758736