

Original Paper

A Multi-Dimensional Framework and Practical Approach for Leveling Chinese Texts for Primary School Students

Xiaoli Dong¹

¹ Xi'an FanYi University, Xi'an, Shaanxi, China

Received: June 08, 2025

Accepted: July 12, 2025

Online Published: July 31, 2025

doi:10.22158/wjeh.v7n4p39

URL: <http://dx.doi.org/10.22158/wjeh.v7n4p38>

Funding Project

Xi'an FanYi University 2025 Institutional Research Project "A Study on the Leveling Chinese Reading Texts for Primary School Students Based on Academic Quality Standards" (Project Number: 2025B10)

Abstract

Text difficulty is not only influenced by linguistic complexity but is also closely related to factors such as readers' cognitive load and cultural context. Current readability formulas have limitations, including overreliance on surface-level text features and neglect of reader differences. Therefore, this paper proposes that when leveling Chinese texts for primary school students, four dimensions should be considered: language, content, cultural context, and genre structure. By combining natural language processing technology to extract features from multiple dimensions such as vocabulary, sentence structure, content, and structure, a text difficulty norm applicable to different grade levels can be established. Through a combination of expert evaluation and data comparison, the difficulty level of the text can be accurately determined.

Keywords

primary school Chinese language, reading texts, difficulty grading, multi-dimensional framework

1. Introduction

Reading is a fundamental way for humans to obtain information and develop thinking skills, and improving primary school students' reading abilities is a current focus of Chinese language education. However, the current extracurricular reading market faces the issue of "rapid growth in quantity but chaotic grading", making it challenging for teachers and parents to select appropriate reading materials. Materials that are too difficult or too easy can affect children's reading interest and sense of efficacy (Stenner et al., 2006), and the core of graded reading lies in accurately assessing text difficulty.

Overseas research on text difficulty grading began earlier, with representative tools such as the Flesch-Kincaid (Flesch, 1948) and Gunning Fog Index (Crossley et al., 2019). Domestic scholars have primarily borrowed methods for assessing the difficulty of English texts when evaluating the difficulty of Chinese texts, focusing on language features such as word difficulty and sentence length and their impact on text reading. However, they have not yet addressed the influence of other text features and individual student differences, significantly reducing the validity of readability formulas. Therefore, this paper proposes that the grading of Chinese reading text difficulty in primary schools should consider language, content, cultural context, and genre structure, and attempts to propose a practical path for text grading.

2. Research on the Readability of Chinese Texts

In recent years, research on Chinese readability formulas has gradually unfolded. Researchers mainly quantify language difficulty through linguistic features such as word diversity and sentence length to infer the difficulty readers encounter when reading Chinese texts. For example, Sun Hanyin (1992) argued that factors such as the number of strokes in characters, the proportion of difficult words, and sentence length can effectively predict the difficulty of Chinese texts. Based on an examination of middle school students' reading processes, he proposed a readability formula. Wu Siyuan et al. (2020) argued that readability formulas should include a multidimensional system of Chinese text features. They used support vector machine algorithms to construct a readability model for Chinese textbooks; Liu Miaomiao et al. (2021) extracted multi-dimensional text features from a corpus of primary school Chinese language textbooks and found that character types, average word difficulty, and the ratio of function words are the primary predictive factors. However, these readability formula studies have the following obvious shortcomings: First, they insufficiently consider the individualized needs of primary school students. Students of different grades have significant differences in language ability and cognitive development, which inevitably affects their understanding of reading texts. Second, current readability formulas primarily examine linguistic features, neglecting key factors such as rhetorical devices, content difficulty, and cultural context. Yet these factors are precisely the important variables influencing the reading comprehension process. This highlights the need for a more comprehensive framework as a reference for grading the difficulty of primary Chinese reading materials.

3. Text Difficulty and Influencing Factors

Text difficulty reflects the cognitive effort expended by readers when reading a specific text (Dale & Chall, 1949). The language, content, genre structure, and cultural context of the text interact with individual reader factors to influence the reading comprehension process (Crossley et al., 2023). primary school students' reading abilities mature gradually with age, and there is significant individual variation. Therefore, when grading text difficulty, it is essential to fully consider the primary roles of language, content, genre structure, and cultural context.

3.1 Language Complexity

Language complexity primarily encompasses the difficulty of language at the word, sentence, and text levels. Scholars primarily use vocabulary complexity, syntactic complexity, and the complexity of cohesive devices to quantify this (Bulté & Housen, 2012). For Chinese reading texts, the lexical level involves the number of strokes in characters, usage frequency, the number of idioms, the proportion of difficult words, and the density of content words; the syntactic level primarily includes sentence complexity, sentence length, and sentence hierarchical structure; the textual level primarily examines the semantic relationships between sentences and paragraphs. In primary school Chinese reading, these factors directly influence the level of text comprehension.

3.2 Content Difficulty

Text content difficulty refers to the complexity of the themes, plot, background knowledge, and implicit information in the text. It not only includes the difficulty of the language used in the text but also the cultural background, historical knowledge, social values, and emotional depth of the text that readers need to understand (Kintsch, 1998). For example, Chinese language texts for lower primary grades primarily consist of fairy tales and stories, featuring simple plots and relatively straightforward language. In contrast, reading texts for upper grades increasingly incorporate abstract concepts and complex plots, with language expressions containing more implicit information, hints, or symbolic expressions. This requires readers to possess relevant background knowledge, reasoning, and comprehension skills.

3.3 Genre and Structure

Genre refers to the type of text, and texts of different genres exhibit significant differences in language use and paragraph structure. For example, narrative texts are based on plot development, have clear structures, and are easy to understand. In contrast, expository texts focus on explaining the nature or process of the subject matter, requiring precise language expression, higher complexity, strong logical coherence, and higher demands on readers' background knowledge and language proficiency. Structure refers to the presentation of information such as paragraph arrangement, sentence connections, and logical hierarchy within a text (McNamara et al., 2014). Presentation methods that align with students' general understanding are easier to comprehend, while those that deviate from such understanding increase difficulty. For example, chronological order is relatively easier to understand than flashbacks or flash-forwards. Therefore, genre and structure are important factors influencing text difficulty.

3.4 Cultural Context

Cultural context includes social and cultural background, customs, values, and other factors required for text comprehension. When selecting texts from Western countries, it is essential to consider Chinese primary students' familiarity with these cultural context elements. Additionally, idioms, proverbs, and other colloquial expressions are closely tied to specific cultural contexts and require background knowledge for full comprehension. Therefore, text reading for different grade levels in primary school must consider the comprehension challenges posed by cultural context.

3.5 Cognitive Characteristics of Primary School Students

The selection of Chinese reading materials for primary school students must take into account their cognitive characteristics. First, younger students have a limited vocabulary and short attention spans, making it easier for them to understand simple and concrete words and phrases. Longer or more complex texts are likely to distract them, so reading texts should be concise and of moderate length. Second, as students progress through the grades, their abstract thinking abilities develop, and their comprehension and attention spans improve. Therefore, the language complexity and content difficulty of reading texts can gradually increase, and the genres of reading texts should also become more diverse. Third, primary school students are emotionally rich, and texts with strong appeal can effectively stimulate their interest in reading.

4. Chinese Language Text Leveling Framework for Primary School

The following text divides the primary school stage into three phases: lower grades (1st–2nd grade), middle grades (3rd–4th grade), and upper grades (5th–6th grade). Based on the cognitive development characteristics of primary school students and the primary factors influencing text difficulty, assessment criteria and difficulty requirements are proposed for each grade level across multiple dimensions, including language complexity (vocabulary, sentence structure), content, text structure, and cultural context.

4.1 Vocabulary Dimension

Reading texts for lower grades should primarily consist of basic everyday vocabulary, with a focus on monosyllabic or common disyllabic words. In middle grades, common polysyllabic words, some idioms, and proverbs should be gradually introduced to help expand vocabulary. By upper grades, the learning of complex words increases, so reading texts should gradually introduce complex vocabulary according to the curriculum guidelines, such as social and historical background terms, more difficult idioms, or cultural proper nouns, to cultivate text comprehension skills.

4.2 Sentence Dimension

According to the cognitive development characteristics of primary school students, simple sentences are the primary syntactic structure of reading texts. Therefore, when grading texts, Chinese language texts containing complex sentence structures such as compound sentences or complex sentences should not be placed in lower grades. The introduction of such complex syntactic structures should be gradual, so reading texts for middle grades can appropriately increase sentence complexity. For upper-grade reading texts, sentence complexity should be increased by flexibly combining complex structures such as simple sentences and compound sentences, and integrating content from different genres to cultivate their comprehension abilities.

4.3 Content and Context Dimension

Regarding the content of reading texts, text grading should consider the progressive nature of content and context. The content of lower-grade texts should primarily draw from familiar everyday scenarios

in children's lives, such as school, home, and amusement parks; with simple language, relatively straightforward plots, and straightforward paragraph structures. For middle-grade texts, gradually introduce stories with complex plots, biographies, science popularization, etc., to better align with the content learned in textbooks. By the upper grades of primary school, students need to read more challenging texts, so the content themes should be diverse, such as social phenomena, historical stories, friendships, scientific experiments, etc. This progressive content design helps foster cognitive development and reading skills in primary school students.

4.4 Text Structure and Discourse Layout Dimension

The structure of reading texts for lower grades should be simple, with few and short paragraphs, and explicit transitional devices should be used to ensure smooth and natural transitions, helping students understand the main content of the text. For middle grades, reading texts can consider using more complex discourse structures, such as parallelism and contrast, to add depth to the content. For upper-grade students, more complex paragraph structures can be designed, using flexible hierarchical structures within the same paragraph, such as simultaneously presenting contrast and transition content, to cultivate their comprehensive analytical thinking abilities.

4.5 Cultural Context and Background Knowledge Dimension

The content of reading texts for lower grades should not only focus on familiar cultural contexts such as students' family life, personal hobbies, school activities, fairy tales, and stories, but also appropriately incorporate elements of excellent traditional Chinese culture, such as the customs of festivals like the Spring Festival, Dragon Boat Festival, and Mid-Autumn Festival, to implement the spirit of documents such as the Ministry of Education's "Guidelines for Incorporating Excellent Traditional Chinese Culture into Primary and Secondary School Curricula and Textbooks" and "Guidelines for Incorporating Revolutionary Traditions into Primary and Secondary School Curricula and Textbooks." Reading texts for middle-grade students can increase the proportion of content related to excellent traditional Chinese culture and also include revolutionary culture, such as revolutionary figures and stories, to lay the foundation for reading in upper grades. Primary school upper-grade students are in a critical phase of rapid development in reading ability. Teachers and parents should guide students to begin exploring various types of books involving more complex cultural and historical contexts, such as historical classics, novels, and texts on scientific inventions, with a focus on cultivating students' ability to recognize and reflect on culture.

5. Steps for Grading the Difficulty of Chinese Language texts in Primary School

Researchers can use natural language processing tools to extract text features across the aforementioned dimensions, establish text difficulty norms, and then compare the proposed graded texts with the norms through expert review to achieve effective leveling of Chinese reading texts, ensuring they align with the cognitive development

5.1 Text Difficulty Analysis

First, use natural language processing tools to analyze language difficulty features, such as word

frequency, proportion of difficult words, sentence length, structural complexity (e.g., simple sentences, compound sentences, etc.), sentence hierarchy, grammatical dependency relationships, and syntactic tree structures (Jin et al., 2020). Second, utilize deep learning models and large language models to analyze the semantic content of the text and assess the alignment between the background knowledge referenced in the text and the cognitive capabilities of primary school students. It is important to note that texts for lower grades should focus on simple everyday stories, while texts for upper grades can gradually incorporate more complex content themes to better cultivate students' higher-order cognitive abilities. Third, the structural and paragraph layout difficulty of the text can be analyzed using methods such as topic modeling to identify how the text's themes are developed. In Chinese reading texts for lower grades, paragraph structures should be concise and clear, with natural transitions; texts for higher grades can gradually incorporate complex paragraph layouts and implicit connections. Finally, by using large language models to analyze cultural contexts, national regions, and background knowledge such as historical figures and place names in the text, the cultural complexity can be assessed to determine whether these elements require students to possess specific background knowledge.

5.2 Text Leveling Path

Based on the above dimensions, the practice of leveling Chinese texts for primary school students can be carried out in the following steps: First, construct a text difficulty norm suitable for different grades. To this end, representative texts from the primary school textbooks of the Ministry of Education, covering various genres such as narrative and expository texts, should be selected to construct a small corpus. Second, utilize natural language processing technology to extract difficulty features from the texts in the corpus based on the aforementioned four dimensions. In terms of linguistic difficulty, focus on analyzing word difficulty, such as stroke count, frequency, and diversity, as well as sentence structure complexity, such as sentence length, proportion of simple and complex sentences, and sentence structure patterns; in terms of content difficulty, use deep learning models and large language models to assess the semantic depth, plot complexity, and background knowledge of Chinese reading texts for different grade levels; In terms of text structure and paragraph layout difficulty, we use topic analysis to quantify the structural hierarchy of the text, identify its logical relationships, and analyze paragraph structure, parallelism, and contrast structures; in terms of cultural context, we use large language models to identify cultural, historical, and geographical elements in the text and assess whether specific cultural background knowledge is required to understand its deeper meaning. Third, based on the multi-dimensional data extracted from texts across different grade levels, calculate the average text difficulty across the four dimensions for each grade level (low, medium, and high), and establish a difficulty norm for texts across different grade levels. For texts awaiting leveling, the same natural language processing tools are first used to extract their feature data, including language difficulty, content difficulty, structural layout complexity, and cultural context depth. These feature data are then compared with the established leveling norms to determine the grade level to which the text belongs. Fourth, experts or teachers are invited to comprehensively assess the difficulty of the texts

awaiting leveling based on their teaching experience and students' cognitive ability levels, and confirm the difficulty level.

6. Conclusion

In summary, this paper analyzes factors such as the cognitive characteristics of primary school students, language features, text content, genre structure, and cultural background, and proposes a preliminary leveling framework for primary school Chinese reading materials. Researchers can use this framework, which includes four aspects—language difficulty, content difficulty, genre and text structure, and cultural context—to level texts. Specific difficulty assessments can be made using natural language processing tools, deep learning models, large language models, and the cognitive judgments of experts or experienced teachers. This framework provides theoretical support and practical pathways for the selection and adaptation of reading texts at the primary school level, aiming to enhance students' reading interest and comprehension abilities through reasonable leveling within an appropriate difficulty range, thereby promoting the comprehensive development of language skills and critical thinking.

References

- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & F. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-46). John Benjamins. <https://doi.org/10.1075/lllt.32.02bul>
- Crossley, S., Heintz, A., Choi, J. S., Batchelor, J., Karimi, M., & Malatinszky, A. (2023). A large-scaled corpus for assessing text readability. *Behavior Research Methods*, 55, 491-507. <https://doi.org/10.3758/s13428-022-01802-x>
- Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, 26, 19-26.
- Flesch, R. F. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221-233. <https://doi.org/10.1037/h0057532>
- Jin, T., Lu, X., & Ni, J. (2020). Syntactic complexity in adapted teaching materials: Differences among grade levels and implications for benchmarking. *The Modern Language Journal*, (1). <https://doi.org/10.1111/modl.12622>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511894664>
- Liu, M. M., Li, Y., Wang, X. M. et al. (2021). An Exploratory Study on Graded Reading: A Study on the Readability Formula of Chinese Based on Primary School Textbooks. *Language and Character Application*, (2), 116-126.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. New York, NY: Cambridge University Press.
- Stenner, A. J., Burdick, H., Sanford, K., & Burdick, D. (2006). The Lexile framework for reading.

Educational Measurement: Issues and Practice, 25(1), 24-37.

Sun, H. Y. (1992). *A Formula for Chinese Readability* (Master's Thesis). Beijing Normal University.

Wu, S. Y., Yu, D., & Jiang, X. (2020). Construction and Validity Verification of a Readability Feature System for Chinese Texts. *World Chinese Teaching*, (1).