

The Confusion of Pinyin in Younger Children and Its Causes

Xinyu Wu*, Xiaoxuan Zheng, Yixin Wang, Wenting Yu, Yu Dai

College of Teacher Education, Quzhou University, Quzhou, Zhejiang, China

Corresponding Author: Xinyu Wu (wuxinyu_456@qq.com)

Abstract: This study selected 74 first and second-grade students from primary schools in Quzhou City as samples. Using written tests, error content analysis, and semi-structured interviews, we analyzed the characteristics and causes of pinyin confusion in young learners. As the core component of Chinese language education in early primary grades, pinyin serves as a fundamental tool for children's character recognition, reading, and Mandarin learning. In actual teaching practice, pinyin confusion frequently occurs and tends to recur, posing a major challenge in early-grade pinyin instruction. The research identified three main types of pinyin confusion: homophone confusion, phonetic similarity confusion, and tone similarity confusion. These issues involve various aspects including initial consonants, vowels, whole syllable recognition, and syllable spelling. Phonetic similarity confusion was the most prevalent. The causes of these three types differ, corresponding to multiple factors such as cognitive development, mechanical memorization patterns, and insufficient teaching training. The conclusions of this study can guide the optimization of pinyin teaching in early grades and provide practical references for addressing pinyin confusion challenges.

Keywords: Young children; Pinyin confusion; Characteristics; Causes; Primary school Chinese teaching

1. Introduction

Pinyin serves as the cornerstone of elementary Chinese language education, acting as a vital tool for young learners to master character recognition, reading, and Mandarin expression. The Compulsory Education Chinese Curriculum Standards (2022 Edition) explicitly requires that lower-grade students 'master Chinese Pinyin, accurately pronounce initial consonants, vowels, tones, and whole syllables, correctly spell syllables, and write initial consonants, vowels, and syllables' [1-3]. As an abstract phonetic system, Pinyin demands that children develop both visual discrimination skills and phonetic perception abilities. Given that young learners are in a transitional developmental stage with predominantly visual thinking patterns, they naturally encounter cognitive challenges when encountering abstract symbols.

When teaching Chinese to lower-grade students, frontline educators frequently report that pinyin confusion remains a persistent issue. Despite repeated classroom interventions—including guided pronunciation drills, error correction, and reinforcement exercises—students often repeat mistakes. As pinyin instruction is among the most challenging components in elementary Chinese education, this study synthesizes existing research findings and examines teachers' practical experiences. It demonstrates that pinyin confusion cannot be simply attributed to student

'carelessness.' The errors exhibit distinct patterns and are influenced by multiple factors. By analyzing the external characteristics of pinyin confusion in young learners and exploring its root causes, this research aims to inform targeted improvements in pinyin teaching methodologies.

1.2 Research Objectives and Significance

This study focuses on three core objectives. First, it collects error samples from written pinyin tests of first and second-grade students, categorizes all pinyin confusion types, and analyzes the distribution patterns across different dimensions including initial consonants, vowels, syllable spelling, and tone. Second, it explores the underlying causes of various pinyin confusions through multiple perspectives such as children's cognitive development, mechanical memorization patterns, and teaching methodologies. Third, by integrating the cognitive characteristics of younger students with the identified patterns and causes, the study proposes targeted and actionable strategies for pinyin instruction optimization, providing practical guidance for frontline teachers to address students' pinyin confusion issues.

Theoretical Significance: By systematically categorizing phonetic confusion characteristics, comparing grade-level variations, and analyzing etiological factors across multiple dimensions, this study enriches theoretical frameworks for elementary Chinese phonics instruction while supplementing existing research on young children's Chinese phonetic acquisition. Current studies predominantly focus on empirical summaries of phonics teaching, whereas systematic analysis of phonetic confusion's comprehensive features and underlying causes remains largely unexplored. This research provides valuable reference directions for future academic investigations in this field.

Practical Significance: This study identifies weak points in phonetic learning among younger students, clarifies the causes of confusion in phonetic content, and highlights developmental differences across grade levels. Teachers can use these findings to guide instructional adjustments. By integrating the confusion patterns and grade-level variations identified in this research, educators can pinpoint learning challenges at each stage. This enables them to optimize teaching methods such as pronunciation demonstrations, comparative instruction, and feedback correction. Furthermore, it allows for designing phonetic activities tailored to the cognitive abilities of younger learners, thereby enhancing classroom efficiency. These efforts help students avoid entrenched pronunciation errors and writing habits, laying a solid phonetic foundation for their future Chinese language studies.

1.3 Study Subjects and Core Conceptual Definition

This study conducted research with first and second-grade students from primary schools in Quzhou City, with a total of 74 participants. The distribution was 38 first-graders and 36 second-graders, including 40 boys and 34 girls. The sample included students with varying academic levels and diverse dialect backgrounds, ensuring representativeness. During the process, incomplete responses were excluded, yielding 72 valid samples: 37 from first grade and 35 from second grade.

The core concepts are defined as follows:

Firstly, the lower grades. The lower grades mentioned in this study refer to first and second graders in primary school. During this stage, children begin systematic learning of Chinese pinyin and establish their own phonetic framework, marking a critical developmental phase. This period also plays a vital role in forming children's phonetic perception and solidifying pronunciation behaviors, making it an indispensable developmental window. Secondly, pinyin confusion. When children complete pinyin recognition, writing, and spelling tasks, they may incorrectly associate two

or more pinyin symbols or syllables with similar appearances or pronunciations, leading to misuse. Based on frontline teaching observations and existing research, this study categorizes pinyin confusion into three types: The first type involves visual similarity confusion, where pinyin symbols closely resemble each other, causing confusion. Examples include b and p, p and q, f and t, m and n. The second type involves phonetic similarity confusion, where pinyin pronunciation positions are close with minimal differences in articulation methods, resulting in auditory and pronunciation confusion. Examples include zh and z, in and ing, n and l. The third type involves tone similarity confusion, where pinyin tones share similar pitch values and contours, causing confusion in distinguishing tones. Examples include second and third tones, second and fourth tones.

2. Literature Review

The teaching of Chinese pinyin is a key field of primary school Chinese education research.

Within existing research on Pinyin instruction, the core positioning of Chinese Pinyin teaching is as a tool for character recognition and a foundation for Mandarin learning. Instruction must focus on its instrumental nature, avoiding excessive theoretical or knowledge-intensive approaches, while aligning with the cognitive development patterns of young children. This constitutes the fundamental principle for Pinyin teaching. Scholars have proposed that Pinyin instruction should adhere to the cognitive patterns of younger students, steering clear of overly abstract or theoretical methods. Instead, it should integrate Pinyin learning with character recognition, reading, and oral expression, enabling the development of Pinyin skills within authentic language contexts. The academic community widely emphasizes that the core objective of Pinyin teaching is to help students build a solid Pinyin foundation without excessive burden, meeting both their current developmental needs and future sustainable learning requirements [4]. This perspective defines the direction and boundaries for Pinyin instruction.

When analyzing the types and characteristics of pinyin confusion, existing research has identified common errors among lower primary school students. First graders frequently experience initial consonant confusion, vowel confusion, and overall recognition of syllable confusion. Students often mix up similar pinyin when writing or spelling [5]. Writing errors predominantly involve misusing mirror symbols like b, d, p, and q, while recognition errors mainly stem from phonetic proximity confusion, which occurs more frequently than other types. Studies on students with pinyin learning difficulties reveal that over 60% of errors involve confusion between level and retroflex sounds, front and back nasal sounds, and lateral nasal sounds. The highest proportion of confusion occurs in these categories. Researchers have identified that tone confusion among lower-grade students primarily affects the second and third tones, stemming from students' lack of concrete perception of pitch variations and unstable pronunciation control. These errors tend to persist stubbornly and do not naturally disappear with advancing grades. Published master's theses indicate that pinyin errors in lower grades span initial consonants, vowels, syllable spelling, and tones. Beyond single-symbol confusion, other errors include missing middle initials in three-syllable words, overall recognition of syllables, and compound vowel confusion. The confusion patterns differ between first and second graders, reflecting grade-specific characteristics.

In studies examining the causes of pinyin confusion, scholars have explored multiple dimensions. Regarding intrinsic factors in children, developmental psychology indicates that visual discrimination and auditory perception abilities in 6-8-year-olds remain in developmental stages. At this age, children's capacity to discern subtle differences in symbols and phonetic nuances is not yet fully

mature, forming the physiological cognitive basis for pinyin confusion in younger students. Research in child language development confirms that ages 6-8 represent a critical period for developing phonetic awareness. During this phase, children's ability to precisely distinguish Chinese phonetic elements continues to evolve rapidly but has not yet reached maturity, making them prone to phonetic confusion. Moreover, Chinese children only achieve rapid development in phonetic awareness after receiving pinyin instruction [6]. Field teaching interviews reveal that lower-grade students often rely on mechanical memorization when learning pinyin tables, failing to establish one-to-one correspondence between pinyin forms and sounds. This fundamental gap is a major contributing factor to the high frequency of pinyin confusion.

Research on linguistic environment reveals that dialects exert negative transfer effects on Mandarin phonetic acquisition. Certain phonetic oppositions—such as the flat-tilted and retroflex oppositions, or front-back nasal oppositions—are absent in dialects. These oppositions impair learners' ability to distinguish corresponding sounds. Scholars argue that teaching pinyin in dialect regions proves more challenging than in Mandarin-speaking areas, as dialectal phonetic habits directly interfere with students' perception and imitation of standard pinyin. This creates a key environmental factor contributing to frequent phonetic confusion.

Research on teaching factors indicates that some teachers overuse phonetics terminology when teaching pinyin, such as 'frontal tip articulation,' 'tongue position elevation,' and 'soft palate elevation.' These technical terms exceed the cognitive level of young children, making it difficult for students to translate abstract concepts into concrete pronunciation actions. Current Chinese pinyin instruction typically focuses on isolated letter and syllable learning, with phonetic forms and pronunciation taught in isolation, lacking coherence [7]. Insufficient comparative teaching for easily confused pinyin characters, lack of integrated reading-writing practice, and a teaching model dominated by mechanical memorization directly exacerbate pinyin confusion, constituting core teaching issues.

A systematic review of existing research reveals that Chinese academia has accumulated substantial studies addressing challenges in pinyin instruction, including common confusion patterns and their underlying causes. However, most current research focuses narrowly on single dimensions, with limited exploration of developmental trends across grade levels, comprehensive characteristics of pinyin confusion, or interactions among influencing factors. These gaps warrant further investigation, and existing empirical data could be expanded. This study integrates prior research by collecting written test samples from first and second graders and semi-structured interviews with teachers and students. By analyzing comprehensive features of pinyin confusion, comparing grade-level variations, and examining the root causes of different confusion types, we aim to fill critical gaps in current scholarship.

3. Research Design

3.1 Research Methods

This study primarily employed the following four research methods:

First, the written test method: After completing the Chinese Pinyin consolidation learning for the current academic year, 74 first and second-grade students from primary schools in Quzhou City were administered a Pinyin test. The test covered four modules: initial consonant table dictation, final vowel table dictation, syllable spelling, and tone marking. All students' Pinyin error samples were collected to provide empirical data for feature analysis and grade-level comparison.

Second, the error content analysis method: Researchers processed pinyin error samples collected from written tests, completing encoding, classification, and statistical analysis. The analytical framework was structured as ‘shape similarity confusion—phonetic similarity confusion—tone similarity confusion,’ with the analysis process incorporating four dimensions: initial consonants, vowels, syllables, and tones. The study ultimately identified distribution characteristics of different types of pinyin confusion and revealed grade-level variation patterns between first and second graders.

Third, the semi-structured interview method was employed. Four Chinese language teachers with over three years of teaching experience and eight first and second graders exhibiting typical pinyin confusion issues were selected for interviews. The student interviews focused on cognitive challenges and learning difficulties related to pinyin confusion, while the teacher interviews addressed teaching difficulties in pinyin instruction, variations in confusion patterns across grade levels, underlying causes, and potential teaching improvements. This approach aimed to clarify the actual prevalence of pinyin confusion in classroom settings and the subjective perceptions of involved parties.

Fourth, literature research method, sorting out the primary school Chinese phonetic teaching, children's speech development, language acquisition field related works, journal papers, master's thesis, for the research of sorting out the concept boundary, determine the research framework, analyze the problem to form logic, adjust the teaching direction to build a theoretical foundation.

3.2 Research Process

The research methodology comprises four distinct phases. Phase I involves preparatory work: compiling existing literature in the field, defining research questions, constructing a research framework, and developing phonetic assessment instruments including error collection forms and semi-structured interview outlines. Phase II focuses on data collection through distributing test materials at local primary schools in Quzhou City, conducting student and teacher interviews, compiling test results, and transcribing interview transcripts. Phase III entails data analysis: categorizing and coding error samples, quantifying error frequency across different types, comparing phonetic confusion patterns between first and second graders, and synthesizing interview insights with theoretical frameworks to identify underlying causes. Phase IV concludes with research synthesis: formulating conclusions based on data analysis, proposing optimized phonetic teaching strategies aligned with current educational needs, and completing the full academic paper.

4. Result Analysis

This study distributed 74 pinyin test papers, with 72 valid responses collected. The analysis identified 789 valid pinyin error samples, including 452 errors in first-grade and 337 in second-grade students. The data was categorized into four modules: initial consonants, vowels, syllable spelling, and tones. The research also examined three types of confusion: homophonic errors, phonetic errors, and tonal errors. The final statistical results are presented below.

4.1 Feature Analysis of Initial Consonant Confusion

The test results of the analysis of the initial consonant module, the errors of the students can be classified into six categories, the specific distribution of each category is organized into Table 1.

Table 1: Distribution of Problems in Initial Sound Learning.

Open question	Number of people who encountered errors	Percentage of valid samples
Incomplete transcription of initial consonants	10	14%
sequence error	12	17%
Confusion between initial consonants and final vowels	6	8%
Homophonic confusion	13	18%
Incorrect initial consonant spelling	13	18%
Incorrect initial consonant format	11	15%

Analysis of specific characteristics and grade-level differences reveals the following conclusions: First, errors in initial consonant memorization show grade-level variations. In first graders, 22% miss initial consonants entirely while 24% make sequence errors, compared to 5% and 8% in second graders. As students progress through grades, their ability to memorize initial consonant sequences improves. Both error types share common patterns: frequently overlooked consonants include y, w, g, k, h, and r, while sequence errors predominantly occur at z, c, s, zh, ch, sh, and r positions, mostly involving reversed order.

Secondly, homophonic confusion constitutes the core issue in initial consonant writing. First graders exhibit a higher error rate than second graders, with both homophonic confusion and initial consonant writing errors accounting for 18% – the highest proportion among all initial consonant errors. Homophonic confusion predominantly occurs between the four consonants b, d, p, and q, with mirror-image confusion between b and d being the most frequent. Writing errors often involve mirror-image issues in pinyin direction, such as: writing the curved hook of t and f in reverse (right-curved strokes incorrectly written as left-curved), writing j's curved hook in reverse (left-curved strokes incorrectly written as right-curved), and writing z, c, and s in reverse directions – all typical examples of homophonic confusion. Error data shows grade-level variations: 62% of initial consonant errors in first grade involve homophonic confusion, while second graders account for 35%. As students develop visual perception skills, this type of error gradually diminishes.

Third, persistent issues with conceptual confusion and formatting standards were observed in first and second-grade assignments. Notably, 8% of students wrote rhymes instead of initial consonants when completing the initial consonant dictation task, indicating unclear understanding of the conceptual boundaries between these phonetic categories. Additionally, 15% of students made formatting errors: for example, b and t were incorrectly placed in the middle upper grid, while f and j were mistakenly written in the middle upper grid despite their proper placement in the upper middle grid. This demonstrates students' inadequate mastery of the standard four-line three-grid pinyin writing system. The error rates showed no significant difference between first and second grades.

4.2 Feature Analysis of Rhyme Mother Confusion

The results of the final syllable module test show that the error rate of students is higher than that of the initial syllable module, and the errors can be divided into four categories, which vary by grade.

First, memory-based errors and sequence errors predominantly occur in first grade, with noticeable changes in second grade. During first-grade dictation tests, 30% of students failed to

complete the syllable finals, while 27% made sequence errors. These percentages drop to 11% and 9% respectively in second grade. Students frequently miss syllable finals such as *ie*, *ue*, *er*, *un*, *un*, *ang*, *eng*, *ing*, and *ong*, primarily involving both front and back nasal finals. Sequence errors mainly involve disordered arrangements of *an*, *en*, *in*, *un*, *un*, and *ie*, *ue*, *er*. Younger students find memorizing the syllable finals more challenging than the initial consonants.

Secondly, phonetic confusion is a core category of vowel errors with inherent persistence. While showing no statistical differences between first and second grades, 42% of vowel errors involve front-back nasal confusion. Typical cases include miswriting 'in' as 'ing' or 'en' as 'eng' in pinyin transcription. For example, 'peng' (朋) is often misspelled as 'pen', and 'qing' (情) as 'qin'. Complex vowel confusion follows at 28%, such as writing 'feng' as 'fong' or 'ting' as 'tin'. 15% of errors involve lateral nasal vowel confusion, mirroring the n-l consonant position confusion. Grade-specific data reveals phonetic confusion accounts for 51% of vowel errors in first grade and 48% in second grade, showing no decline. This phenomenon results from the negative transfer effect of dialectal environments in western Zhejiang and the developmental stage-specific characteristics of students' auditory discrimination abilities, which does not naturally improve with higher grade levels.

Thirdly, conceptual confusion is prevalent in Chinese language education, particularly between complex vowels and whole syllables. This issue is evident in first and second-grade instruction, where students frequently miswrite 'un' as 'yun' or 'ing' as 'ying' when copying vowel charts. Such errors reveal that students lack a clear understanding of the conceptual distinctions and pronunciation rules between these two categories. Therefore, targeted teaching should be implemented to clarify these distinctions.

4.3 Feature Analysis of Syllable Spelling Errors

The analysis of the syllable spelling module's test results shows that students' errors can be categorized into three main types, which demonstrate a developmental pattern corresponding to grade levels.

First, incomplete three-syllable structures are the most common spelling errors in syllable formation. Students often overlook the medial vowel between initial consonants and final vowels, resulting in incorrect syllable combinations like writing 'shuo' as 'sho', 'xiong' as 'xong', or 'nuo' as 'nou'. These errors reveal students' insufficient mastery of three-syllable structure rules and frequent omissions during spelling practice. Error patterns vary across grade levels: such errors account for 47% of total syllable errors in first graders, while the proportion drops to 32% in second graders. This decline indicates that increased exposure to spelling exercises enhances students' understanding of three-syllable structures.

Secondly, the most prevalent spelling errors occur in first grade, where students frequently confuse initial consonants with vowels, resulting in syllable formations that violate pinyin rules. Typical examples include misreading 'you' (from 'you') as 'iou', 'ye' (from 'ye') as 'yie', and 'shui' (from 'shui') as 'shiu'. These errors primarily stem from students' lack of clear understanding of vowel structure and the principles of whole-syllable recognition. By second grade, however, the incidence of such errors shows a noticeable decline.

Third, the transfer of initial and final consonant confusion persists throughout the first and second grades of elementary school. When students spell syllables, they tend to transfer the confusion between similar-looking and phonetically similar single characters to the spelling of complete syllables. For example, when annotating the pinyin for 'shou' (hand), they might write 'sou' instead

of 'shou', or for 'chong' (insect), they might write 'cong' instead of 'chong'. This phenomenon mirrors the confusion characteristic of initial and final consonant modules. The issue of single-character confusion directly impacts the accuracy of complete syllable spelling, and the transfer of phonetically similar character confusion remains evident in second grade.

4.4 Feature Analysis of Tone Confusion

The test results of the tone annotation module show that students' errors fall into three categories:

First, tone proximity confusion remains the most prevalent error type, showing persistent persistence. The confusion between the second and third tones accounts for 53% of all tone errors, followed by the second and fourth tones at 31%. These findings align with existing research: the second tone is a rising middle tone while the third is a falling rising tone, both sharing similar tonal contours. Younger students struggle to accurately distinguish subtle pitch variations, which explains the high concentration of proximity confusion. Grade-level data shows a decline: first graders account for 58% of tone errors, while second graders show 47%, though proximity confusion still dominates as the most common error type.

Secondly, marking light tones correctly poses a significant challenge for younger students. Many struggle to distinguish light tones from other tones, often mistaking them for the first tone. For example, they may label 'ya' in 'ya' as the first tone, or confuse light tones with the fourth tone in 'men'. Some even misplace the 'zi' at the end of words like 'shoes' or 'chairs' —a tone that should be light—marking it as either the first or fourth tone. These errors reveal students' weak grasp of light tone pronunciation rules and labeling standards, with no noticeable difference in error rates between first and second graders.

Thirdly, incorrect tone marking positions are prevalent among first-grade students, with their frequency decreasing after entering second grade. First-graders exhibit higher rates of tone marking errors that violate the rules, commonly involving incorrect letter assignments for the vowels ai, ei, and ui. The root cause of these issues lies in students' weak retention of tone marking rules and insufficient practice to reinforce them. In second-grade students, however, the proportion of such errors shows a decline.

4.5 Analysis of Interview Results

Through compiling student interview data, we identified three persistent cognitive challenges among those with typical pinyin confusion: First, most students recognize the visual similarity of pinyin initials (b, d, p, q) but struggle to discern their subtle distinctions. Second, even when aware of differences, they often fail to recall the precise characteristics that distinguish them—particularly the orientation of the semicircles (b, d, p, q), which poses the greatest memory barrier. Third, while students can fluently recite the initial and final syllable lists, they cannot immediately produce correct pronunciations when presented with a single pinyin, failing to establish the phonetic-morphemic correspondence. This mechanical memorization approach indicates students have not truly grasped the semantic significance embedded in pinyin symbols.

Through organizing teacher interviews, frontline educators identified three core causes of pinyin confusion. First, students habitually memorize pinyin through mechanical repetition without practicing out-of-order recognition or integrating phonetic-semantic matching in reading-writing exercises. Second, as pinyin is an abstract linguistic symbol, lower-grade students struggle to grasp it

due to a lack of concrete associative memory support. Third, teachers rarely provide comparative explanations for easily confused pinyin in daily instruction, nor do they conduct targeted error analysis or remedial exercises. Field teachers reported that first-graders frequently confuse similar-looking pinyin characters and have weak memory retention of pinyin rules, while second-graders mainly struggle with phonetic similarity confusion and oversight in mastering light tones and tones. Teaching strategies should be tailored to address specific challenges at different grade levels.

5. Discussion on Causes

This study discusses the causes of three kinds of pinyin confusion separately, by integrating the results of the written test error statistics, the grade difference comparison data, the teacher-student interview records and the existing theoretical research contents.

5.1 Causes of Homophony

First, children's visual perception development exhibits stage-specific limitations. At six years old, spatial perception remains underdeveloped, lacking a stable orientation concept[6]. Their ability to distinguish graphic details and spatial directions is not yet mature, with significant improvement occurring around age eight—a finding consistent with the grade-level differences observed in this study. For example, first-graders exhibit higher rates of shape similarity confusion compared to second-graders. Pinyin symbols, primarily composed of simple strokes, show only minor differences in orientation (e.g., b-d, p-q) or stroke direction (e.g., f-t, j). These subtle distinctions are often overlooked by younger children, leading to shape similarity confusion and mirror-image writing errors. This is largely attributed to their underdeveloped spatial perception, which results in unstable perception of left-right and up-down orientations, contributing to the high prevalence of mirror-image type similarity confusion.

Secondly, children exhibit deficiencies in intentional attention and observational skills. During the early elementary years, their attention span remains relatively short, and their intentional attention capacity is underdeveloped. When performing tasks like rapid writing or reading, their focus tends to scatter, making it difficult to meticulously capture the details of pinyin symbols. This often leads to errors such as confusing similar-looking pinyin. Student interviews reveal that most students report 'not noticing the differences between similar pinyin', indicating that they have not yet developed the habit of actively observing and comparing details. This subjective factor is a key reason why similar pinyin confusion persists.

Third, the lack of concrete memory support is evident. Pinyin symbols exhibit abstract morphological differences, while young children primarily rely on visual thinking. During instruction, educators fail to connect abstract symbol variations with tangible objects or actions, hindering students from forming stable memories. In student interviews, most reported difficulty remembering the orientation of each pinyin's semicircle. This stems from the absence of concrete memory strategies, as relying solely on rote memorization prevents students from firmly grasping subtle differences between symbols.

5.2 Causes of Homophonic Confusion

First, the negative transfer effect in dialect environments. Language transfer theory suggests that learners' native dialect phonological systems influence standard Mandarin acquisition. When dialects

lack phonetic opposition to Mandarin, negative transfer occurs—learners struggle to distinguish corresponding sounds. Our study sample from Quzhou City in western Zhejiang Province demonstrates poor differentiation between flat and retroflex sounds, as well as front and back nasal sounds in the local dialect. Students raised in dialect environments lack daily exposure to standard Mandarin phonetic input, hindering the development of stable phonemic opposition perception. This leads to persistent phonetic confusion. The study reveals that such confusion does not naturally diminish with academic progression, confirming that dialect environments exert lasting effects on phonetic acquisition.

Secondly, children's auditory perception and articulatory development are limited. Younger children's auditory and speech organs are still developing, unable to distinguish subtle differences in frequency and pitch, nor can they accurately differentiate between similar-sounding sounds like flat and retroflex consonants or front and back nasal sounds. Their control over articulatory organs is not yet fully mature, making it difficult for them to discern phonetic variations or precisely regulate the movements of the tongue tip, tongue position, and airflow. As a result, they cannot execute the corresponding articulatory actions, leading to phonetic confusion.

Third, there are shortcomings in teachers' instructional methods. Interview records reveal two issues in phonetic teaching practices. First, excessive use of specialized phonetics terminology like 'tongue touching the upper gum', 'backward tongue retraction', and 'soft palate elevation' exceeds young learners' cognitive capacity. Students struggle to grasp these abstract concepts and translate them into concrete pronunciation actions. Second, there's a lack of comparative teaching. For easily confused initial consonants and vowels, teachers neither organize systematic content nor demonstrate comparisons through multisensory methods. Instead, they rely solely on repetitive reading drills, which prevents students from establishing clear phonetic distinctions and developing stable discrimination skills. Furthermore, some teachers provide delayed feedback on pronunciation errors, focusing only on correcting sounds without explaining the root causes. This approach allows incorrect pronunciation habits to solidify, making it harder to correct phonetic confusion.

5.3 Causes of Near-Confusion

First, the concept of tone is inherently abstract. In Mandarin, tones are formed through pitch variations, with the core being pitch control. Younger children primarily rely on visual thinking and struggle to grasp abstract concepts like the 'five-degree scale' or 'pitch values'. They also find it challenging to translate abstract pitch changes into concrete pronunciation control. For example, the second tone is a rising middle tone, while the third is a falling rising tone. Although their pitch contours are similar, the differences lie only in the starting and turning points, with minimal variation. Younger students cannot accurately perceive these nuances or mimic them, often leading to confusion between similar tones.

Secondly, children's auditory perception development exhibits inherent limitations. Their ability to discern subtle pitch variations develops progressively with age. Younger students' perception of tones remains at a holistic stage, unable to precisely differentiate nuanced pitch changes. Field teachers' feedback during interviews consistently indicates this developmental pattern: students effortlessly distinguish between the first and fourth tones, but struggle with the second and third tones or the second and fourth tones. This developmental stage aligns perfectly with the maturation of children's auditory perception capabilities.

Third, the teaching process lacks concrete design. Some teachers, when teaching tone patterns,

merely repeat reading and require students to memorize rhyming tone songs without using concrete multisensory teaching tools. This prevents students from associating tones with tangible objects, making it difficult to establish stable tone perception. Additionally, some teachers focus more on teaching initial consonants and final vowels, allocating insufficient effort to tone instruction and failing to arrange systematic training for tone discrimination and pronunciation. The content dedicated to teaching light tones is inadequate, leading to confusion between tones with similar pitch values and persistent errors in marking light tones.

5.4 Causes of Grade Differences in Pinyin Confusion

The findings from this study reveal that phonetic confusion among younger children manifests differently across grade levels. First graders primarily confuse phonetic symbols with similar visual forms, often making errors related to rule-based memorization. Second graders tend to confuse phonetic symbols with similar pronunciations, with mistakes mainly occurring in distinguishing tones and recognizing light tones. These differences stem from both the developmental characteristics of children's cognitive growth and the varying teaching priorities at different educational stages.

In cognitive development, first-grade students aged 6-7 are in a rapid growth phase of visual-spatial perception and intentional attention. They struggle with abstract symbol recognition and frequently make shape-similarity confusion errors. By second grade (ages 7-8), their visual perception improves, and shape-similarity issues naturally resolve. However, negative transfer from dialect environments persists, and limitations in auditory perception remain. Tone-related phonetic confusion emerges as the core challenge.

In terms of teaching methodology, first-grade phonics instruction primarily focuses on new lesson content, emphasizing the recognition, reading, and writing of phonetic symbols, along with foundational phonetic rules and comparative teaching of easily confused sounds, while providing limited specialized training. Second-grade phonics instruction shifts to consolidation and application, lacking systematic specialized review of phonetics. The curriculum also fails to address persistent issues such as phonetic confusion caused by dialectal influences and the rules governing light tones, leaving these stubborn problems unresolved.

6. Conclusion and Recommendations

6.1 Research Conclusions

In the stage of learning pinyin for lower grade children, the confusion of pinyin can be divided into three types, which are the confusion of similar forms, the confusion of similar sounds and the confusion of similar tones. The confusion of similar sounds may occur in the initial consonants, finals, syllables and tones. The confusion of similar sounds occurs more frequently than the other two types and is the most common type of confusion.

Secondly, grade-level differences exist in phonetic confusion among younger children: In Grade 1, errors mainly involve confusion between phonetically similar characters and deviations in memorizing phonetic rules. These errors tend to improve gradually with advancing grades and cognitive development. By Grade 2, the focus shifts to confusion between phonetically similar characters and improper use of tone marks (light tones). These issues are more persistent and do not naturally resolve with higher grades.

Third, the confusion of different types of pinyin stems from various causes. The confusion of similar-looking characters arises from developmental limitations in lower-grade children's visual

perception, intentional attention, and observational skills. The confusion of phonetically similar characters results from dialectal environmental negative transfer, the developmental state of children's auditory and articulatory organs, and the multifaceted influence of teaching methods employed by educators. The confusion of tone-related characters is associated with the abstract nature of the concept of tones, children's developmental level of auditory perception, and insufficient concrete teaching support in classrooms.

Fourth, teachers' instructional methods significantly impact children's phonetic confusion. When teaching relies heavily on abstract terminology, lacks sufficient comparative exercises for easily confused pinyin, fails to adjust content according to different learning stages, provides imprecise feedback and correction, or underinvests in teaching tone and soft sounds, it directly exacerbates children's phonetic learning confusion and solidifies incorrect pronunciation habits.

6.2 Suggestions for Teaching Optimization

(1) Targeted Teaching Strategies by Educational Stage

The pinyin instruction for first graders focuses on two key aspects: preventing confusion between homophonic symbols and mastering fundamental application rules. Teachers can employ visual aids like color-coded symbols, directional comparisons, and gesture simulations to help students better perceive subtle differences in pinyin. Interactive activities such as 'Clap Games' and 'Find Friends Games' [8] encourage students to develop active observation and comparative skills. Basic elements like standardized syllable structures and tone marking should be taught concurrently with reading and writing exercises, helping students establish a 'shape-sound' correspondence in pinyin and moving beyond rote memorization.

In teaching pinyin to second-grade students, educators should prioritize addressing persistent issues like homophonic confusion and tone-light tone errors. By incorporating dialect features from western Zhejiang, teachers can design targeted comparative exercises focusing on common phonetic mix-ups among local students, such as flat and retroflex sounds, as well as front and back nasal sounds. These exercises should employ multisensory methods—combining lip movements, auditory comparisons, and tactile experiences—to minimize the negative transfer effects of dialects. Specialized lessons on tones and light tones can be developed using concrete tools like gestures, role-playing, and rhythmic nursery rhymes to help students establish stable tone perception. Pinyin training should be integrated with character recognition and reading instruction, allowing students to reinforce their skills in real-life contexts rather than through isolated drills.

(2) Type-specific Intervention Strategies

First, to address homophonic confusion in pinyin, educators can incorporate engaging teaching aids during instruction. For instance, using letter cards or handcrafted colorful visuals [7] helps students distinguish similar-looking pinyin characters, fostering differentiated memory. Patterns like 'right lower circle bbb, left lower circle ddd, right upper circle ppp, left upper circle qq' can also be utilized. Additionally, tiered comparative exercises can be designed—starting with individual character analysis, progressing to syllable recognition, and culminating in sentence-level application—thereby progressively enhancing students' discernment skills.

Secondly, to address homophonic confusion, educators may implement a 'perception-contrast-consolidation' multisensory teaching model. Instead of overusing abstract phonetic terminology, teachers should explain pronunciation techniques in child-friendly language, such as 'for flat sounds, press the tongue tip against the teeth; for retroflex sounds, lift the tongue tip.'

By incorporating local dialect features, teachers can compile frequently confused initial consonants and vowels for systematic contrastive listening and reading practice. Additionally, increasing immersive Mandarin exposure through activities like reading texts with pinyin, nursery rhymes, and stories can help mitigate the negative transfer effects of dialects.

Third, to address tone confusion and soft tone errors, teachers can design concrete, scenario-based teaching activities. For instance, the 'Five-Step Gesture Tone Adjustment Method' can be employed, where gestures demonstrate pitch variations to help students visualize abstract tonal concepts. Activities like 'Tone Matching' and 'Contextual Reading' enable students to master soft tone pronunciation rules through real-word contexts. Implementing a real-time feedback system allows precise correction of tone errors, analysis of their causes, and ultimately helps students develop stable tone control skills.

(3) Strategies for Holistic Teaching Optimization

First, we should adjust the teaching mode which is based on mechanical memorization, implement the teaching principle of 'combining form and sound, combining reading and writing, combining learning and using', integrate the learning of pinyin into the training of character recognition, reading and oral expression, so that students can master the pinyin in the actual application and get rid of the learning path of memorizing the pinyin table in isolation.

Secondly, we should pay attention to individual differences among students. Based on the common knowledge confusion types of different students, we should design matching tutoring programs according to students' learning conditions, carry out one-to-one precision error correction training, arrange special exercises of corresponding knowledge modules, and avoid the one-size-fits-all teaching mode.

Third, strengthen home-school collaboration by guiding parents to use standard Mandarin in daily family communication. Provide written notices in pinyin for parents to read aloud at home, and label household furniture and appliances with pinyin [9]. Work with schools to implement pinyin recognition and reading exercises, creating a conducive learning environment. By integrating educational efforts from both home and school, we can jointly address students' pinyin confusion issues.

Acknowledgements

This research was supported by the following projects:

The key cultivation project of the 15th 'Challenge Cup' Undergraduate Entrepreneurship Plan Competition of Quzhou University, 'Chinese Children's Voice - AI Pinyin Learning System for Overseas Chinese Children Enabled by edge computing';

Quzhou University 2025 National College Student Innovation and Entrepreneurship Training Program Project 'International Chinese Education Data Platform: Intelligent Diagnosis and Prediction Model for Multi module Learning Behavior' (Project No. 202511488065);

2025 Zhejiang Province University Student Science and Technology Innovation Activity Plan (New Talent Program) 'Pen Falling Star River: Intelligent Reading of Primary and Secondary School Composition Research on the Design and Practice of Writing Enhancement Platform' (Project No. 2025R450A013);

2024 Quzhou University the construction project of the 'Mentor+Project+Team' Innovation and Entrepreneurship Workshop, 'Smart Book Boat' (Project No. SH20241101).

References

- [1] Cao Jingyun. A Cold Reflection on the Chinese-Style 'Flipped Classroom'. *Invention and Innovation (Educational Informatization)*, 2017, (10):20-24.
- [2] Lü Caiyun. A Brief Discussion on the Teaching Method of Chinese Pinyin for Grade One. *Extracurricular Chinese Language*, 2021, (10):141-142.
- [3] Wu Qingchun. Love 'Striving' to Win, Procurement 'I' Can Do It — Evaluation and Suggestions on the Pinyin Reading Section of Peking University's Chinese Language Teaching. *Primary School Chinese Teaching*, 2025, (18):16-19.
- [4] Xu Yi. Pinyin Teaching Concepts under the Unified Textbook System. *Basic Education Curriculum*, 2020, (21):50-56.
- [5] Sun Lili. A Study on the Current Status of Pinyin Learning Among First-Grade Students. *Jiangsu University*, 2023. DOI: 10.27170/d.cnki.gjsuu. 2023.002086.
- [6] Xu Fen, Dong Qi, Yang Jie, et al. Development of Chinese phonological awareness in primary school children. *Psychological Science*, 2004, (01):18-20. DOI: 10.16719/j.cnki.1671-6981.2004.01.005.
- [7] Yang Xiufang. Research on Problems and Countermeasures in Primary School Chinese Pinyin Teaching. *Yunnan Normal University*, 2014.
- [8] Fu Xiaoli. Exploring the Implementation of Fun Teaching Methods in Primary School Chinese Pinyin Instruction. *Chinese Language Journal*, 2013, (08):171-172.
- [9] Pan Ying. Some Thoughts on Optimizing Pinyin Teaching in Primary Schools. *Science Popularization (Science Education)*, 2013, (03):93. DOI: 10.16728/j.cnki.kxdz. 2013.03.01.