Curriculum and Teaching Methodology (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/curtm.2025.080809
ISSN 2616-2261 Vol. 8 Num. 8

Research on Language Problem Diagnosis and Optimization Schemes in Adolescents' Recitation Training Based on AI

Tiantian Wang[1], Yi Dai[1], Luqi Zhao[2]

[1]City University of Macau, Macau, 999078, China
[2]Wuhan Optics Valley Tangxun Lake School, Wuhan, Hubei, China

Keywords: Artificial Intelligence; Adolescent Recitation; Language Issues

Abstract: Within this analytical framework, recitation emerges as a pivotal method for linguistic expression and comprehensive literacy development, demonstrating unique educational value during adolescents' language acquisition. However, traditional recitation training approaches exhibit limitations such as monotonous formats, delayed feedback, and lack of personalized support, often resulting in pronunciation errors, unstable intonation, superficial semantic comprehension, and weak pragmatic transfer abilities in young learners. To address these challenges, this study proposes constructing a diagnostic and intervention model for language issues in recitation training using artificial intelligence technologies—including speech recognition, semantic analysis, and multimodal interaction systems—while exploring its theoretical significance from an hermeneutic perspective. The model aims to respond to current concentrated attention on typical language barriers in recitation, further developing dynamic feedback mechanisms and systematic training restructuring strategies. Guided by cognitive development theories, these mechanisms and strategies not only align with the inherent characteristics of adolescent language acquisition but also provide theoretical foundations and practical pathways for intelligent language education, charting a clear developmental trajectory.

1. Introduction

As a comprehensive training method, recitation integrates language skills, emotional expression, and cultural understanding, occupying a prominent position in the macro framework of language education. For adolescents in the critical language development phase, the quality of recitation training significantly impacts both the efficiency of language system construction and the maturity of expressive abilities. In this analytical context, particular attention should be paid to the profound influence of teaching methodologies. Currently, recitation training often lacks systematic implementation. Analysis reveals that language issue identification predominantly lacks systematic support, with feedback mechanisms being neither sufficiently persistent nor stable, while training methods often fail to adequately accommodate individual differences. Given the subtlety of these findings, structural imbalances in language development frequently emerge. Considering the multidimensional nature of these evidence, this study demonstrates that constructing an intelligent

training system centered on problem diagnosis-leveraging AI technologies such as speech analysis, semantic modeling, and contextual simulation-holds promise for optimizing recitation education. This represents a crucial step forward within an evolving conceptual framework. Notably, such systems possess inherent potential to address the aforementioned shortcomings.

2. Core Characteristics of Adolescent Recitation Language

2.1 The Developmental Phases of Language Cognition and the Imbalance of Expressive Ability

During language development, adolescents typically tend to master pronunciation and sentence patterns earlier than emotional and semantic expression, tending to point toward what appears to be a phenomenon that might be characterized as ‘form outweighing content’ in recitation. For instance, in reciting '*Farewell to Cambridge Again*,' students may ostensibly imitate rhythm and stress accurately but appear to struggle to convey what appears to represent the gentle sadness of parting-their voices often remaining stable yet seemingly lacking emotional depth[1]. What seems to generally emerge from these observations is that adolescents tend to focus on pronunciation accuracy while predominantly overlooking what appears to constitute the integration of emotion and context.

As adolescents mature, their language proficiency demonstrates progressive development. Upper elementary students primarily focus on mimicking standard pronunciation, which often indicates limited mastery of complex sentence structures. During junior high school, while their control over pronunciation and rhythm improves, emotional expression and semantic comprehension remain underdeveloped. By high school, students begin integrating language and emotion, yet their ability to comprehend complex texts and contextual awareness remains superficial. Given these theoretical complexities, training methods must adapt to students 'cognitive development. Even technical perfection may result in emotionally hollow performances without genuine expression. This suggests that recitation training should start from students' current cognitive language levels, gradually strengthening the balance between vocal control and emotional delivery[2].

2.2 The Imitative Tendency in Emotional Intonation Acquisition and the Dilemma of Style Formation

When learning emotional intonation, adolescents typically exhibit a pronounced tendency to imitate. Bandura's social learning theory suggests that individuals acquire language behaviors through imitation of external models. Given the complexity of these theoretical relationships, students in recitation training often mimic teachers or exemplary recordings to learn pitch, rhythm, and speech rate. However, the traditional interpretation faces challenges: this reliance on imitation may result in formulaic emotional expression lacking individuality. Observations reveal that junior high school students can largely replicate intonation patterns, yet their emotional delivery often appears monotonous with limited vocal tension, rarely demonstrating truly distinctive emotional styles [3]. The outcomes of this imitation-based learning approach suggest that formulaic emotional expression lacks intrinsic resonance and contextual understanding-a phenomenon worthy of further exploration.

Within the current analytical framework, existing recitation courses and evaluation standards appear to further suppress stylistic diversity. A prominent issue in many training models is the excessive emphasis on technical metrics like standardized pronunciation and speech rate, which seems to reflect a neglect of emotional variety and individuality. Evidence shows that some schools over-rely on "standardized model readings," which appears to limit students' creative space.

Therefore, to address this problem, this model suggests that curriculum design should emphasize tasks rich in emotion and contextuality, encouraging students to adjust intonation and emotional expression in different situations, transitioning from mere imitation to creative expression.

3. Diagnosis of Linguistic Problems in Current Recitation Training

3.1 Technical Deficiencies in Pronunciation Accuracy and Intonation Control

Adolescents frequently encounter pronunciation inaccuracies and unstable intonation during recitation. While accurate pronunciation serves as the fundamental criterion for evaluating recitation quality, research indicates that teenagers often exhibit deviations in initial consonants, vowels, and tones. In Mandarin, phonemes like "zh" and "ch" are particularly prone to confusion, while vowel combinations such as "an" and "ang" frequently show errors. Analysis of acoustic features-including fundamental frequency (pitch variation), intensity (volume fluctuation), and rhythm (articulation duration)-reveals these pronunciation deviations. Studies demonstrate that pitch variations typically reflect tonal modulation, while intensity changes correlate with stress placement and intonation processing [4]. Evidence shows that many junior high school students struggle with pitch control, especially during transitions between high and low syllables. This phenomenon challenges traditional interpretation methods, making it difficult for them to precisely convey nuanced emotional expressions.

In this context, inadequate intonation control also leads to issues such as inaccurate tone conversion and unstable speech rate, which should not be overlooked. At the end of sentences, tones often exhibit delays or deviations-a phenomenon worthy of deeper exploration, as it hinders the expression of emotional layers and progression. For instance, when reciting "Spring River Moonlit Night," some students failed to properly utilize rising tones and pauses in the line "The river winds around fragrant meadows, the moon illuminates flower groves like snowflakes" (this is an example of classical poetry for easier understanding; the original English translation "The moon climbs high, the river and city wide" is a literal rendering). This resulted in incomplete emotional expression. Given the complexity of these theoretical relationships, it appears that educators could greatly benefit from using quantifiable indicators such as "pitch range," "stress position," and "rhythmic stability" within this broader analytical framework for evaluation and analysis.

3.2 Dual Barriers in Semantic Comprehension and Pragmatic Transfer Ability

Adolescents often face challenges in semantic comprehension and pragmatic transfer during recitation. Research indicates that many students focus primarily on literal meanings while lacking deeper textual understanding. The discourse analysis framework provides a valuable approach to evaluate students' grasp of implicit meanings and emotional depth-a finding particularly noteworthy [5]. When reciting classical poetry, students may pronounce words correctly but struggle to convey the emotions embedded in the verses. For instance, when reciting "Thoughts on a Tranquil Night," students can fluently recite the line "Before my bed the moonlight glows," yet often fail to convey the loneliness and homesickness conveyed through the poem, resulting in superficial understanding of the work.

From this specific interpretive perspective, adolescents often lack sufficient flexibility in adjusting intonation across different genres when practicing pragmatic transfer. Reciting poetry and prose typically requires content-specific modulation of tone and rhythm-poetry demands emotional variation while prose generally requires fluidity and calmness. Given these methodological considerations, many students appear generally challenged in making such adjustments. For instance, when reciting the narrative prose "My Dream," they often adopt overly agitated intonation,

evidently overlooking the quiet, contemplative nature inherent in prose. Notably, this deficiency in pragmatic transfer ability suggests weakened expressive adaptability among students, further confirming reduced congruence between emotions and context [6]. Therefore, this pattern indicates that training should focus on enhancing students' pragmatic transfer capabilities across diverse texts to achieve more effective language expression.

4. AI-Supported Optimization Pathways for Recitation Language

4.1 Precision Diagnosis and Personalized Feedback System Based on Speech Analysis

The speech analysis system based on deep speech recognition and acoustic modeling demonstrates significant potential for delivering highly accurate pronunciation diagnosis and personalized feedback. Within this broader analytical framework, the system extracts acoustic features such as pitch, intensity, duration, and rhythm to perform real-time analysis of students' recitation performance. Given the complexity of these theoretical relationships, the technology integrates phoneme recognition, intonation matching, and stress detection functions, enabling seemingly precise identification of pronunciation issues like phoneme errors or intonation deviations while providing targeted correction suggestions [7]. For instance, pitch analysis appears to offer a method for evaluating the consistency between intonation and emotional expression. Notably, experimental data reveals that the system achieves approximately 92% accuracy in phoneme recognition, with speech rate and intonation feedback delays controlled within 1 second, indicating guaranteed real-time precision.

In educational applications, considering the subtleties of these findings, personalized feedback systems appear to enable customized training paths based on students 'performance. These systems typically record each student's voice characteristics and scores to build individual profiles, thereby guiding targeted training. For instance, evidence shows that when detecting accent deviations during recitation of "Spring River Moonlit Night," the system recommends accent control exercises. This pattern suggests that after four weeks of training, students 'accent accuracy improved by approximately 15%. In this context, it's equally noteworthy to explore how such precise feedback mechanisms significantly enhance teaching efficiency and help students continuously improve their recitation skills through personalized guidance.

4.2 Construction of a Multimodal Contextual Training Platform and Reconstruction of Pragmatic Competence

Recitation training, particularly those incorporating speech diagnosis and personalized feedback, generally indicates the need to enhance situational comprehension and pragmatic transfer capabilities. Within this broader analytical framework, multimodal situational training platforms primarily focus on the three-dimensional collaboration of "language-image-emotion." By integrating natural language processing, emotion recognition, and situational modeling technologies, these systems appear to establish dynamic interactive training mechanisms. Given the intricate theoretical relationships, the platform typically manifests as a four-module architecture: semantic understanding, emotion recognition, situational generation, and adaptive feedback. The semantic module analyzes textual meaning, the emotion module identifies emotional characteristics in speech, the situational module reconstructs realistic scenarios through multimodal fusion, while the adaptive feedback module optimizes speech rate, intonation, and expressive strategies in real-time.

In practical applications, considering the subtleties of these findings, when students recite "Farewell to Cambridge Again", the system appears to automatically generate images and music matching the poem's poetic atmosphere, seemingly guiding them to experience the gentle emotions

of parting. Evidence shows that when the system detects monotonous tones or repetitive rhythms, it typically provides immediate prompts like "slow down your speech" or "use a softer tone" to encourage active adjustments. When reciting "My Dream", the platform generally allows switching between different contextual backgrounds, primarily enabling students to freely transition between narration, emotional expression, and reflection. This gradually cultivates sensitivity to contextual changes while enhancing natural language control capabilities. Particularly noteworthy is the system's real-time adaptive guidance capability [8].

This analysis reveals that the platform operates through a "contextual simulation-behavioral feedback-transfer reconfiguration" training framework. It stimulates situational awareness through multimodal inputs, enhances self-correction via real-time feedback, and restructures language strategies across text types through pragmatic transfer. The system advances speech technology toward pragmatic competence, boosting emotional expression and linguistic adaptability in real-world scenarios. Furthermore, it provides a practical technical and methodological framework for intelligent recitation education.

5. Conclusion

This study appears to utilize artificial intelligence as the core technology, aiming to construct a language problem diagnosis and optimization system designed for adolescent recitation training. The system’s architecture is its functioning through speech recognition, semantic analysis, and multimodal context modeling, tending to provide what appears to be predominantly full-chain support, from pronunciation correction to pragmatic transfer. What the system appears to reveal, within this broader analytical framework, is its capacity to largely address the limitations of traditional recitation training, given the complexity of these theoretical relationships, such as delayed feedback and what appears to represent a lack of contextualization. What this tends to indicate is the study's proposal of what seems to constitute an operable intelligent intervention model that tends to enable substantially quantifiable and traceable improvement in recitation teaching, appearing to provide evidence that may support apparently significant educational value. What these findings seem to point toward, in light of these methodological considerations, is the system’s potential for typically wide applicability in school recitation courses, oral expression training, and online language learning platforms, tending to point toward what appears to be improved pronunciation accuracy and naturalness of expression in real contexts. What appears to emerge from this evidence, however, is that the research scope and corpus types still appear to tend to be somewhat limited, and what seemingly demonstrates a need for further testing is the system’s adaptability across age groups and dialects. What appears to follow from this analysis, therefore, is that future studies can explore pragmatic transfer modeling, personalized feedback for emotional expression, and expansion to multilingual environments within these evolving conceptual parameters, to continuously refine the theory and practice of intelligent recitation training.

References

[1] Feng, J. (2025). How artificial intelligence empowers elementary students’ recitation and expression. Reading, Writing and Arithmetic, (30), 55–57.(In Chinese)

[2] Jiao, Q. (2020). A study on the representation of paralinguistics in broadcasting and hosting recitation expression-Based on recitation guidance and teaching practice. Audiovisual, (11), 129–131.(In Chinese)

[3] Xie, R. (2020). A preliminary study on strengthening narrow paralinguistic training in middle school Chinese oral communication teaching (Master’s thesis, Shanghai Normal University). Master 2020.000118.(In Chinese)

[4] Xu, H. (2023). Phonetics for Broadcasting and Hosting [M]. Communication University of China Press, 214.(In Chinese)

[5] Nie, H. (2020). Guidance on Recitation and Storytelling [M]. Nanjing University Press, 251.(In Chinese)

[6] Gu, X. (2013). Recitation: Let children's hearts leap with vitality. Examination Weekly, (64), 38.(In Chinese)
[7] Wang, D. (2012). Children's Language Arts Course [M]. Central Compilation and Translation Press, 253.(In Chinese)
[8] Zhang, S. (2007). Poetry Recitation [M]. Communication University of China Press, 281.(In Chinese)