Academic Journal of Education
ISSN: 3080-9940 (Print), ISSN: 3080-9959 (Online) | Vol. 2, No. 2, 2026

A Praat-Based Study of Phonological Acquisition among Junior High School Students

Xiaoyan Zhan [1], Yan He [2]

[1] School of Foreign Languages, Huanggang Normal University, Huanggang, Hubei, 438000, China
[2] Qihuang Middle School, Huanggang, Hubei, 438000, China

Abstract: The study investigates the status-quo of English phonological acquisition among Chinese junior-high-school students. Fifty-six Grade-8 students from Qihuang Middle School (Huanggang) participated in this experiment. Reading excerpts were drawn from Unit 7 Session B 2b of the PEP Grade-8 Volume 2 textbook. Speech recordings were captured using Cool Edit Pro 2.1, four-level corpus annotation was performed in Praat, and quantitative statistical analyses were conducted via Excel. Visual phonetic analyses were implemented across four dimensions: word stress, sentence stress, rising-falling intonation, and pronunciation accuracy. Results reveal notable phonological biases concerning tone-group segmentation, stress placement, and intonation patterns. Such deviations primarily stem from L1 negative transfer, constrained pedagogical approaches, and insufficient corrective feedback. Accordingly, this paper proposes pedagogical interventions underpinned by speech-visualization techniques. The findings offer empirical evidence and practical implications for refining junior-high-school English phonetics instruction, particularly for teaching adapted to computer-scored man-machine dialogue assessments.

Keywords: Praat; Speech Visualization; Junior English; Phonological Acquisition; Pronunciation Teaching.

1. Introduction

According to the *Compulsory Education English Curriculum Standards (2022 Edition)*, phonology is intrinsically linked to semantics, and language realizes its social communicative functions through phonetic forms. The curriculum requires junior-high-school learners to recognize the educational value of phonetics as well as its meaning-conveying functions in contextual settings. Nevertheless, junior-high-school English phonetics teaching in China faces considerable practical challenges. Constrained by the long-standing tradition of prioritizing written proficiency over oral competence, many teachers attach inadequate importance to phonetics instruction and allocate disproportionate class time to examination-oriented knowledge such as vocabulary and grammar. Phonetic practice is often fragmented and mechanized, with insufficient contextualized training activities. Discontinuities between primary- and secondary-school phonetics curricula further widen gaps in learners' phonetic foundations, which in turn impedes students' vocabulary retention, listening comprehension, and oral communicative competence. Driven by educational informatization, intelligent speech assessment has been adopted for junior-high-school graduation examinations across over 40 provincial-level administrative regions in China. The emergence of computer-assisted man-machine dialogue assessments imposes higher requirements on learners' pronunciation accuracy and oral reading fluency. Within conventional classrooms, both teachers and learners struggle to perceive abstract phonetic attributes directly, and corrective feedback is frequently delivered without objective acoustic evidence. Speech-visualization techniques offer a promising solution for pedagogical innovation. Tools such as Praat convert auditory phonetic features into visual spectrograms, enabling a shift from experience-driven phonetics instruction toward evidence-based practice and mitigating inherent limitations of traditional teaching.

The study adopts Praat-based visual acoustic analysis to investigate phonetic characteristics of Chinese junior-high-school English learners and identify typical pronunciation biases together with their underlying causes. It intends to equip front-line educators with practical Praat-analytic protocols and instructional examples so as to enhance the scientific rigor and relevance of phonetics teaching. Furthermore, this research develops age-appropriate visualization-oriented pedagogical strategies to foster learners' self-monitoring and self-correction capacities and optimize phonetic learning outcomes. Finally, it generates empirical evidence for pedagogical adjustments responding to computer-scored speaking assessments, facilitating constructive alignment between examination evaluation and classroom-based English instruction.

2. Literature Review

2.1. Second-Language Phonological Acquisition Research

Research on second-language (L2) phonological acquisition has long centered on learners' pronunciation deviations and their underlying sources. In a study investigating word-stress acquisition among Chinese EFL learners, Bu (2021) identified three major contributors to so-called "Chinese-accented" English: misallocated lexical stress, insufficient amplitude contrast between stressed and unstressed syllables, and pervasive sentence-level over-stress. These findings highlight persistent suprasegmental challenges for Chinese L2 English learners [1]. Drawing on Wells' (2006) intonation framework [2], Bu (2016) further reported functional deficiencies in Chinese learners' intonation performance, including inappropriate tone-group chunking, misaligned nuclear-accent placement, and inappropriate pitch-pattern selection, which impair spoken fluency and pragmatic competence [3]. L1 negative transfer constitutes a major source of L2 phonological difficulty. Chen (2019) observed that Chinese EFL learners tend to produce

multiple stress peaks, read at slower tempos, and employ limited prosodic adjustments in connected speech — patterns attributed to the syllable-timed rhythmic properties of Mandarin[4]. Shang and Bu (2009) corroborated L1 interference in vowel production, lending empirical support to L1 transfer accounts within L2 phonological acquisition[5].

2.2. Pedagogical Applications of Speech-Visualization Technology

Speech-visualization techniques transform auditory phonetic input into visual representations to facilitate learners' perceptual awareness of subtle phonetic features. Zhuang and Bu (2011) argued that visualization tools help resolve the well-known pedagogical dilemma of phonetic phenomena that are "sensible yet hard to articulate", and are particularly well-suited for suprasegmental phonology instruction[6]. Early scholarship largely focused on software demonstrations; for instance, Zhuang and Bu (2011) illustrated instructional practices using BetterAccent Tutor for suprasegmental training [7]. In recent years, Praat has gained widespread research-pedagogical acceptance owing to its open-access availability and comprehensive analytical functions. Shang (2016) explored Praat's utility in EFL classrooms, noting that its visual outputs foster learners' capacity for self-monitoring and pronunciation remediation[8]. Chen (2013) adapted Praat for Mandarin pronunciation training, whose operational workflows offer valuable insights for EFL phonetics pedagogy[9]. While prior work validates the feasibility of visualization-assisted teaching, much of the existing literature remains descriptive and tool-oriented, with few systematic classroom-based empirical investigations.

2.3. Research on Junior-High-School EFL Phonetics Instruction

Junior high school represents a formative period for establishing stable L2 pronunciation habits; nevertheless, phonetics pedagogy at this level suffers from notable deficiencies. Gao and Zhang (2018) contended that insufficient instructional emphasis and monotonous teaching approaches constrain Chinese adolescents' overall language development[10]. Aligned with curriculum guidelines, Zhou and Liu (2022) advocated embedding phonetics practice within unit-thematic contexts so that phonetic forms are taught in conjunction with semantic and situational meaning[11]. Regarding assessment practices, Li (2021) proposed a "rank-progression evaluation framework" for pronunciation teaching[12]. This procedural model structures phonetic remediation through successive stages: diagnosing phonetic deficits, correcting mispronunciations, and consolidating accurate articulatory patterns. Against the backdrop of computer-scored speaking assessments, Li and Li and Wang (2020) underscored the importance of targeted phonetic diagnosis and championed technology-enabled precision instruction[13]. Zhang (2024) recommended integrating thematic-meaning inquiry into phonetics lessons to construct a scaffolded learning sequence of perception-comprehension-appreciation-application, thereby enhancing learners' overall phonetic literacy[14].

In summary, extant studies have documented the instructional promise of speech visualization and catalogued recurrent pronunciation difficulties among Chinese EFL learners. Still, critical gaps remain. First, most visualization-related research targets university-level learners, whereas empirical evidence for junior-high-school populations remains scarce. Second, many studies isolate individual phonetic features without adopting multi-dimensional systematic analysis. Third, proposed pedagogical strategies are predominantly theoretically conceived and lack grounding in empirical classroom data. Addressing these gaps, the present study develops a multi-dimensional visual-acoustic analytical framework, systematically examines junior-high-school learners' phonological difficulties, and puts forward evidence-informed instructional recommendations.

3. Research Questions

Adopting an experimental approach, this study draws on Praat-enabled speech visualization, alongside subjective auditory judgement and acoustic measurements, to profile the phonological characteristics of Chinese junior-high-school EFL learners and investigate their L2 phonological acquisition and oral production performance. The specific research questions are formulated as follows:

1.In light of the new national curriculum standards and the demands of computer-scored speaking assessments, how can Praat be deployed to conduct multi-faceted visual phonetic analyses covering word stress, sentence-level stress, rising-falling intonation, and pronunciation accuracy?

2.What are the salient phonological features manifested in junior-high-school EFL learners' oral production, and what underlying factors give rise to their pronunciation difficulties?

3.How can the findings derived from visual-acoustic analysis and the identified sources of phonological difficulties inform the optimization of junior-high-school EFL phonetics teaching strategies?

4. Research Design

4.1. Participants

The participants were 56 Grade-8 EFL learners (30 female, 26 male) with a mean age of 14 years and approximately four years of formal English learning experience. All participants took part in the experiment on a voluntary basis. None had received specialized pronunciation training, and their learning backgrounds were representative of general junior-high-school student populations. Grade-8 learners were selected because this developmental phase constitutes a critical period for establishing L2 pronunciation habits. At this stage, students possess foundational English competence and are capable of completing the assigned reading-aloud task.

4.2. Materials

The reading-aloud stimulus was excerpted from Unit 7 Session B 2b in Volume 2 of the PEP Grade-8 English textbook. The passage contained 304 words centering on giant-panda conservation. The theme was relatable to adolescent learners and matched participants' proficiency level, with an average sentence length of 15 words. Target lexical items had been pre-introduced in course materials. Using textbook-sourced reading material guaranteed stimulus validity and enabled direct pedagogical transfer of the research outcomes. Received Pronunciation (RP) audio from the official textbook CD served as the pronunciation reference. One week prior to data collection, participants were given access to the text for preview, eliminating lexical comprehension barriers and securing valid speech data.

4.3. Instruments

Speech signals were recorded using Cool Edit Pro 2.1 with a sampling rate of 44 100 Hz, 16-bit resolution, and mono-channel mode. A Takstar PCM-5550 microphone was positioned 30 cm in front of each participant. Recordings were conducted in a sound-attenuated language laboratory where background noise was kept below 30 dB. Praat (Version 6.2.09) was employed for phonetic annotation and acoustic analysis. Microsoft Excel 2019 was used for quantitative processing, calculating feature frequencies and duration distributions to perform descriptive and difference-based statistical analyses.

4.4. Data Collection and Annotation

Prior to formal recording, participants listened to the reference audio and engaged in 5 minutes of imitative practice. Each participant entered the recording booth individually and read the passage at a natural speaking rate. Re-recording was permitted for misread segments, with a maximum of three attempts per participant. Audio files were labeled following the naming convention: Student ID-Gender-Date. All recordings were annotated in Praat via four-tier Text Grid objects. The orthographic tier contained the reference transcript. The oral-performance transcription tier captured actual spoken output and pauses, with a pause threshold set at 100 ms. The boundary tier marked tone-group boundaries according to Cruttenden's criteria[15]. The prominence tier annotated nucleus and stress locations within each tone group; nuclear accents were determined by combined acoustic evidence of pitch, duration, and intensity.

4.5. Data Analysis

Following annotation, data were extracted across four analytical dimensions: (a) word stress, including accuracy rates and error categories; (b) sentence-level stress, calculating the proportion of stressed words and making comparisons against the RP reference pattern; (c) rising-falling intonation, computing intonation accuracy for distinct sentence types and categorizing pitch-pattern errors; (d) pronunciation accuracy, measuring vowel- and consonant-correctness rates and assessing the appropriateness of pause placement and pause duration. Data were compiled in Excel. Descriptive statistics summarized global phonological profiles. Frequency analysis identified recurrent pronunciation errors, and qualitative case-based analysis explored underlying causal factors.

5. Results and Analysis

5.1. Analysis of Lexical-Stress Patterns

Lexical stress constitutes a fundamental component of the English prosodic system and exerts critical influences on lexical meaning and speech-stream rhythm. In the present study, acoustic parameters including fundamental frequency (F0), intensity, and duration were extracted via Praat to identify stress locations in participants' oral production, which were subsequently compared against the reference Received Pronunciation. The descriptive statistics are presented in Table 1.

Overall lexical-stress accuracy stood at 59.7%. Disyllabic words yielded the highest accuracy, whereas polysyllabic words exhibited the most severe stress-placement errors. This finding aligns with Bu's (2021) observation that stress misplacement represents a pervasive difficulty among Chinese EFL learners [1].

Table 1. Accuracy and Error Distribution of Lexical Stress Production (N = 56)

Stress Category	Correct Items	Incorrect Items	Accuracy Rate (%)	Major Error Type
Disyllabic words	892	348	72.1	Excessive stress on the initial syllable
Trisyllabic words	513	407	55.8	Inadequate stress on the penultimate syllable
Polysyllabic words	186	314	37.2	Erroneous stress placement
Total	**1591**	**1069**	**59.7**	—

Two main error patterns were identified: participants tended to assign stress to word-initial syllables by default, or produced unstable stress locations for polysyllabic items. For instance, 38 participants incorrectly stressed the initial syllable in endangered. Acoustic measurements from Praat revealed that the first syllable in these mispronunciations displayed significantly higher fundamental frequency, intensity and duration relative to the RP reference production. Such mis-behaviour can be attributed to L1 negative transfer. As Mandarin is a tone language with little perceptual contrast between syllable prominence, learners struggle to detect English lexical-stress distinctions. This supports Chen's (2019) claim that Mandarin rhythmic patterns interfere with L2 English pronunciation [4].

5.2. Analysis of Sentence-Stress Characteristics

Sentence-level stress serves the communicative function of foregrounding informational focus within spoken discourse. The quantitative data demonstrated that participants produced an average of 5.2 stressed words per sentence, which constituted 48.3 % of total word count. By comparison, the RP reference production contained 3.1 stressed words on average, accounting for 29.5 % of all words. The results indicated prominent multi-stress and over-stress tendencies among junior-high-school EFL learners, who frequently stressed both notional words and function words indiscriminately. For example, in the target long sentence, numerous participants assigned comparable acoustic prominence to function words such as to, the, and of as they did to content-bearing notional words. Praat-derived intensity measurements verified that the intensity gap between participants' stressed function words and notional words was minimal. In contrast, function words in the RP reference exhibited markedly lower intensity values. Two contributing factors account for such over-stress behaviour. First, learners tend to prioritize individual-word pronunciation accuracy while neglecting prosodic organization at the discourse level (Zhou & Liu, 2022) [11]. Second, classroom instruction seldom provides context-embedded teaching concerning sentence stress, and teachers offer insufficient explicit

training on discourse prosody.

5.3. Analysis of Intonation Patterns

Intonation constitutes a critical linguistic device for conveying mood and emotion, and it plays an essential role in transmitting pragmatic meaning. The present investigation examined intonation realizations across three core sentence types: declarative sentences, yes-no questions, and wh-questions. Analytical outcomes are summarized in Table 2.

Table 2. Statistics of intonation accuracy rates (N = 56)

Sentence Pattern	Total Sentences	Correct Numbers	Accuracy Rate (%)	Main Error Type
Declarative sentences	328	213	65.0	Sentence-final rising intonation
Yes-no questions	164	95	57.9	Sentence-final falling intonation
Wh-questions	112	89	79.5	Flat, non-variable intonation
Total	604	397	65.7	—

Overall intonation accuracy amounted to 65.7 %. Wh-questions yielded the highest accuracy, whereas yes-no questions obtained the lowest performance. The predominant error for declarative sentences lay in inappropriate sentence-final pitch rise, stemming from L1 negative transfer induced by Mandarin declarative-sentence intonation conventions. For yes-no questions, participants frequently produced sentence-final falling contours. This phenomenon suggests that learners failed to grasp the pragmatic functions of intonation and confused the canonical mapping between sentence types and pitch patterns. Praat-generated fundamental-frequency contours visually demonstrated the divergence in pitch trajectories between learner productions and the RP reference.

5.4. Tone-Group Segmentation and Pause-Feature Analysis

Tone-group segmentation exerts a decisive influence on spoken-discourse fluency and intelligibility. Annotated and analyzed following Cruttenden's criteria, the data revealed that participants segmented each sentence into 3.2 tone groups on average, compared with merely 2.1 tone groups in the RP reference. This indicates excessive fragmentation of tone-group boundaries among learners. Participants' mean pause duration reached 185 ms, with individual pauses extending up to 450 ms, substantially longer than the reference value of 120 ms. Moreover, 38 % of all pauses occurred at non-prosodically-licensed positions, disrupting semantic coherence of utterances. Most learner-produced pauses were respiratory-motivated breaks rather than semantically motivated prosodic-unit pauses. Such behaviour originates from word-by-word reading strategies and insufficient global comprehension of the reading passage. These empirical findings resonate with Gao and Zhang's (2018) argument that prosodic-flow training is largely neglected within junior-high-school EFL instruction [10].

5.5. Pronunciation-Accuracy Analysis

Pronunciation accuracy encompasses both segmental and suprasegmental dimensions. At the segmental level, vowel accuracy stood at 68.3 %, while consonant accuracy reached 75.6 %, identifying vowels as the primary area of learner difficulty. This result corroborates earlier findings reported by Shang and Bu (2009) [5]. Typical vowel-related misarticulations included confusions between /iː/ and /ɪ/, lack of perceptual and production contrast between /æ/ and /e/, and monophthongization of target English diphthongs. Consonant errors mainly manifested as substitution of /θ/ and /ð/, inadequate retroflexion for /r/, and word-final consonant deletion. Such deviations largely arose when learners substituted English-specific phonemes with acoustically similar Mandarin phonetic counterparts. On the suprasegmental plane, participants produced speech at an average rate of 105 words per minute, in contrast to the RP reference rate of 135 words per minute. Learners exhibited slow and highly variable speaking tempo, directly associated with abundant inappropriate pauses, and a prominent word-by-word decoding reading style. This empirical outcome diverges from the observations documented by Liu et al. (2013) [16]. The discrepancy is tentatively attributed to the absence of visualization-based intervention training within the present research design.

6. Discussion

6.1. Summary of Phonological Features among Junior-High-School EFL Learners

Synthesized quantitative and qualitative data reveal five salient phonological characteristics in junior-high-school learners' English oral production. First, learners demonstrate an unstable stress system. Frequent lexical-stress misplacement in polysyllabic words, together with sentence-level over-stress and impoverished light-heavy rhythmic contrasts, align with Bu's (2021) findings and indicate learners' inadequate command of English rhythmic patterns[1]. Second, learners exhibit monotonous intonation. Confused intonation realizations across sentence types and insufficient pitch variation prevent learners from encoding fine-grained semantic nuances via intonation, echoing the functional intonation biases reported by Bu (2016)[3]. Third, tone-group segmentation is highly fragmented. Excessive tone-group counts, ill-located pauses and overly long pause durations degrade overall discourse fluency. Such patterns appear more pronounced among junior-high-school participants than among university-level EFL samples, pointing to developmental-stage-specific phonological competence (Chen, 2019)[4]. Fourth, segmental deviations are pervasive. Learners struggle to articulate English-unique phonemes; recurrent difficulties include monophthongization of diphthongs and word-final-consonant elision, exposing deficits in both segmental perception and production. Fifth, robust L1-negative-transfer effects are observable. Transfer from Mandarin's tone-syllable-timed prosodic system gives rise to blurred stress prominence, flattened intonation and disorganized pausing. These observations lend support to the contrastive-analysis hypothesis within second-language-acquisition research.

6.2. Analysis of Sources of Pronunciation Difficulties

6.2.1. Language Transfer

Mandarin functions as a tone language, where lexical meaning is differentiated by pitch tones. By contrast, English is a stress-timed language whose rhythmic organization is built upon alternations between stressed and unstressed syllables. Substantial systemic discrepancies exist between these two prosodic systems. When learners transfer Mandarin articulatory habits to English production, syllables tend to be produced with roughly equal loudness, lexical stress becomes indistinct, and Mandarin tonal patterns are inappropriately substituted for English intonation. At the segmental level, learners frequently replace language-specific English phonemes with acoustically similar Mandarin counterparts, giving rise to abundant pronunciation deviations.

6.2.2. Pedagogical Limitations

Phonetics instruction suffers from inadequate curricular allocation. In most contexts, phonetic-symbol teaching is confined to the first semester of Grade 7, with limited sustained follow-up practice. Instructional approaches remain monotonous, prioritizing mechanical imitation and repetition over explicit feature instruction and visual demonstrations. Teaching content tends to be fragmented, focusing predominantly on segmental phoneme drills while neglecting suprasegmental prosody. Corrective feedback largely depends on teachers' subjective auditory judgement without objective acoustic evidence. The present study corroborates Zhang's (2024) observation regarding fragmented and mechanized phonetics pedagogy in EFL classrooms[14].

6.2.3. Deficiencies in the Assessment System

Existing assessment criteria are ambiguous and lack operational descriptions of observable phonetic features. Evaluation predominantly relies on teachers' subjective ratings with little technological support. Assessment emphasis falls disproportionately on segmental accuracy, whilst suprasegmental dimensions including stress, intonation and prosodic fluency are under-represented. Formative assessment practices are poorly implemented, so learners receive timely feedback on their phonetic deficits only rarely. Although Li's (2021) "rank-progression evaluation framework" offers promising remedial potential, it has not yet been widely adopted within junior-high-school classroom settings[12].

6.2.4. Inadequate Implementation of Speech-Visualisation Technologies

Real-world deployment of visualization-assisted phonetics teaching faces considerable barriers. Practising teachers often lack hands-on training for tools such as Praat and possess limited technical literacy. Commercially available teaching software mostly targets segmental phoneme practice, with scarce support for suprasegmental visualization. Insufficient school hardware and institutional research resources further create a disconnect between technical afford and everyday classroom practice.

6.3. Pedagogical Strategies Informed by Speech-Visualization Technology

6.3.1. Multi-modal Phonological Instruction

Leveraging Praat's visual-acoustic capabilities, an auditory-visual-kinesthetic multi-modal pedagogical framework can be implemented. Spectrogram and waveform outputs render abstract phonological phenomena perceptible. Learners contrast their own acoustic productions against the RP reference, supplemented by articulatory illustrations to reinforce kin-aesthetic awareness (Yang, 2020)[17]. For practical classroom implementation, instructors utilize fundamental-frequency contours to identify stress peaks, compare pitch trajectories across sentence types to establish intonation-pattern associations, and apply time-annotation tasks to practise appropriate pausing, thereby improving instructional precision.

6.3.2. Thematic-Context-Embedded Phonetics Pedagogy

Consistent with Zhou and Liu's (2022) unit-thematic-context framework [11], phonological practice can be embedded within meaning-driven thematic exploration. Illustrated by the giant-panda-conservation theme, Praat enables lexical-stress analysis, visual-aided prosodic reading, and intonation-guided role-play. Thematic inquiry mitigates instructional fragmentation and actualizes the perception-comprehension-appreciation-application learning cycle (Zhang, 2024) [14]. Rather than remaining decontextualized drills, phonetic activities become meaning-oriented communicative tools, boosting learner motivation and practical linguistic competence.

6.3.3. Optimizing Assessment and Feedback Mechanisms

Drawing upon Li's (2021) rank-progression evaluation framework [12], a visualization-enhanced formative-assessment system can be established. Multi-faceted evaluation criteria address segmental accuracy, stress appropriateness, intonation plausibility, and prosodic fluency. Acoustic metrics extracted via Praat are triangulated with teachers' subjective judgement to integrate quantitative measurements and qualitative appraisal. Visualized progress reports and individual learner portfolios document developmental trajectories. Synchronous in-class feedback and asynchronous automated post-class feedback jointly foster learners' phonological self-monitoring capacities.

6.3.4. Advancing Teacher Professional Development

High-quality phonetics instruction presupposes well-prepared practitioners. Tiered professional training should strengthen teachers' theoretical phonological knowledge, hands-on Praat proficiency, visualization-oriented lesson-design competence, and skills for acoustic-evidence-informed evaluation. Schools ought to build collaborative pedagogical-research communities to support collective curriculum development, case-exchange workshops, and learner-centred action research. Local school-based resource repositories should also be developed to nurture educators equipped with both phonological expertise and digital-pedagogical literacy.

6.3.5. Pedagogical Adaptations for Computer-Scored Speaking Assessments

Computer-administered man-machine dialogue assessments demand targeted instructional adjustments. Praat-supported acoustic comparisons remediate pronunciation errors that negatively impact scoring; rhythm-clapping and metronome-assisted drills enhance oral-reading fluency. Mock assessments familiarise learners with marking rubrics and system properties, whilst learners are trained to carry out pre-examination self-audits using visualization tools (Chen, 2021; Li & Wang, 2020) [18, 13]. Speech-visualisation techniques thus support test preparation alongside the development of authentic communicative performance.

7. Conclusion

The study deployed Praat to acoustically analyse reading-aloud productions from junior-high-school EFL learners and uncovered recurrent phonological challenges: polysyllabic lexical-stress misplacement, sentence-level over-stress, monotonous intonation, excessive tone-group fragmentation, and salient vowel-segment errors. Such difficulties stem from L1 negative transfer, pedagogical limitations, flawed assessment structures, and under-utilization of speech-visualization resources. The five sets of evidence-based strategies—multi-modal instruction, thematic-context embedding, optimized formative assessment, teacher-capacity building, and assessment-oriented pedagogical adaptation—are proposed to enhance junior-high-school phonetics education.

As a cross-sectional investigation, this study is constrained by its single-school, single-class sample, the absence of controlled interventions, and its predominant focus on word- and sentence-level phonology. Future research should expand participant recruitment, conduct longitudinal follow-ups, design pedagogical intervention studies, extend analysis to the discourse level, develop customized instructional tools, deepen technology-pedagogy integration, refine empirically grounded phonological teaching practices, and promote learners' phonological proficiency as well as cross-cultural communicative competence.

Acknowledgments

The work was supported by the Teaching-Research Project of Huanggang Normal University under Grant No. 2022CE67 (Project Title: A Praat-Based Study of English Phonology among Junior High-School Students).

References

[1] Bu, Y. H. (2021). A study of stress acquisition by Chinese EFL learners. In English-Chinese comparison and translation 14 (p. 63). Shanghai Normal University.

[2] Wells, J. C. (2006). English intonation: An introduction. Cambridge University Press.

[3] Bu, Y. H. (2016). A study on intonation acquisition by Chinese EFL learners. Foreign Language Teaching and Research, 48(04), 569-582.

[4] Chen, H. (2019). A visualization-based study on discourse features of college students' English reading aloud. Journal of Southwest University of Science and Technology (Philosophy and Social Sciences Edition), 36(01), 82-88.

[5] Shang, C. Y., & Bu, Y. H. (2009). An analysis of English vowel errors based on experiment. Journal of Tianshui Teachers College, 29(06), 135-137.

[6] Zhuang, M. Q., & Bu, Y. H. (2011). Research on teaching visualization. Research on Audio-Visual Education, 02(02), 92-98.

[7] Zhuang, M. Q., & Bu, Y. H. (2011). BetterAccent Tutor and visual teaching of suprasegmental phonemes. Audiovisual Teaching of Foreign Languages, 02(02), 31-38.

[8] Shang, C. Y. (2016). Application research of visualization software Praat in English class. Journal of Mudanjiang University, 25(04), 148-151.

[9] Chen, Y. (2013). Application of visualization software Praat in Putonghua teaching for normal students. Journal of Henan University of Science and Technology, 33(06), 118-120.

[10] Gao, X. X., & Zhang, Z. X. (2018). On junior high school English teaching. Journal of Heilongjiang Institute of Education, 37(12), 89-91.

[11] Zhou, W., & Liu, C. Y. (2022). English teaching practice in junior high schools in the context of unit themes. English Teaching and Research in Primary and Secondary Schools, 03(03), 16-20.

[12] Li, X. H. (2021). The practical exploration of "Duan system" in junior high school English evaluation. English Teaching and Research in Primary and Secondary Schools, 09(09), 67-70.

[13] Li, X. H., & Wang, M. F. (2020). Diagnosis and treatment: Junior high school English teaching strategies in the context of man-machine dialogue. English Teaching and Research in Primary and Secondary Schools, 01(01), 29-32.

[14] Zhang, Z. F. (2024). Research on junior high school English teaching based on thematic meaning inquiry. English Teaching and Research in Primary and Secondary Schools, 08(08), 31-35.

[15] Cruttenden, A. (2014). Intonation (3rd ed.). Cambridge University Press.

[16] Liu, X. B., Lin, W. H., Zhang, W., et al. (2013). An experimental study on English imitation reading teaching based on visualization. Research on Audio-Visual Education, 34(04), 81-86.

[17] Yang, H. X. (2020). A study on the effectiveness of visualization technology assisting English intonation acquisition. Heilongjiang Education (Theory and Practice), 07(07), 89-90.

[18] Chen, H. (2021). An analysis of junior English teaching strategies in the context of man-machine dialogue. New Curriculum, 29(01), 139.