ISSN: 2669-2341 10.33422/ijsfle.v3i1.708

Driving Digitalization: Action Research to Innovate Pronunciation Pedagogies by Building a Virtual Voice Laboratory

Angel Yim Hung Fan

Centre for English and Additional Languages, Lingnan University, Hong Kong SAR, China

E-mail: fanyh@connect.hku.hk

Cite this article as: Fan, A. Y. H. (2024). Driving Digitalisation: Action Research to Innovate Pronunciation Pedagogies by Building a Virtual Voice Laboratory. *International Journal of Second and Foreign Language Education, 3*(1), 27-55. https://doi.org/10.33422/ijsfle.v3i1.708

Abstract

This paper explores the potential of digitalization in supporting English as a Second Language (ESL) students' development of pronunciation skills through a parallel action research. The author created a Virtual Voice Laboratory (VVL) platform to innovate pronunciation pedagogies in higher education; and simultaneously conducted an ongoing action research study to examine three research questions. The questions examined the efficacy of the VVL in facilitating blended learning, self-directed practice, and self-evaluation within a pronunciation course, and its impact on student engagement, participation, and pronunciation skill development. This VVL as a tailormade digital platform was implemented with undergraduate ESL students in Hong Kong's higher education context. It enabled blended learning, wherein students engaged with the mechanics and articulatory movements essential for improving English pronunciation through self-directed practice using the VVL. Adopting an action research methodology, the study investigated the process of implementing digitalization and blended learning strategies. This was achieved through the analysis of student recordings, sound chart comparisons, and qualitative-quantitative feedback collected during a pilot study. The findings reveal nuances of student experiences, their learning trajectories, and the results from the pilot study, illuminating the impacts and outcomes facilitated by the e-learning approach. The research holds valuable implications for integrating digitalization and blended learning in language education, contributing to the ongoing scholarly dialogue on the transformative potential of technology-mediated approaches in fostering dynamic and impactful learning experiences for diverse student populations.

Keywords: action research, blended learning, innovative pedagogy, e-learning, Teaching English as a Second Language (TESL)

1. Introduction

The COVID-19 pandemic profoundly disrupted traditional models of language education, necessitating a critical rethinking of pedagogical approaches within the higher education landscape (Hodges et al., 2020; Reimers & Schleicher, 2020). As universities worldwide have grappled with the challenges of remote and hybrid learning, there is a pressing need to explore the strategic integration of digitalisation and blended learning strategies to support students' language development (Garrison & Kanuka, 2004; Salmon, 2013). Rooted in Vygotsky's seminal theory of the zone of proximal development (ZPD; Vygotsky, 1978), this imperative is particularly salient for language courses, where learners require tailored support and scaffolding to navigate the complexities of acquiring new linguistic skills. The ZPD emphasises

the crucial role of social interaction and guidance from more capable peers or instructors in facilitating learning (Vygotsky, 1978). In the wake of the pandemic, when face-to-face interactions were severely curtailed, the judicious adoption of digital tools and blended learning approaches offers a promising avenue to recreate dynamic learning environments (Bowen, 2012; Garrison & Vaughan, 2008).

Recent trends in language education further underscore the value of embracing digitalisation and blended learning (Hampel & Stickler, 2005; Sharma & Barrett, 2007). The proliferation of online resources, interactive platforms, and multimedia materials has transformed the landscape of language instruction, enabling more personalised, autonomous, and immersive learning experiences (Kukulska-Hulme & Shield, 2008; Thorne & Payne, 2005). Moreover, the pandemic accelerated the need for flexible, technology-mediated pedagogies that can adapt to evolving circumstances and diverse student needs (Hodges et al., 2020; Rapanta et al., 2020). Against this backdrop, the present study explores the implementation of digitalisation and blended learning strategies within the context of university-level language courses. By situating the research within Vygotsky's theoretical framework (Vygotsky, 1978) and responding to current trends in the field, this paper aims to illuminate the pivotal role of technology-enhanced approaches in supporting students' language development and fostering more dynamic, learner-centred educational experiences (Garrison & Kanuka, 2004; Salmon, 2013).

Cognizant of the pressing need to enhance the teaching and learning efficacies within an elective pronunciation course, the author undertook a strategic intervention grounded in the principles of action research (Mertler, 2017; Stringer, 2013). Recognizing the transformative potential of digital technologies and blended learning approaches (Graham, 2006; Oblinger & Oblinger, 2005), the author developed a specialized platform to facilitate a more dynamic and personalized learning experience for students undertaking the pronunciation elective course. However, existing literature on the integration of technology in language education has primarily focused on the broader impacts on student engagement and learning outcomes, with limited empirical evidence on the specific strategies and approaches for enhancing pronunciation skills (Hew & Cheung, 2014; Wang et al., 2019). The impetus for this initiative stemmed from a keen awareness of the evolving landscape of language education, particularly in the wake of the COVID-19 pandemic (Hodges et al., 2020; Rapanta et al., 2020). The unprecedented disruptions to traditional modes of instruction underscored the urgent need for universities to explore innovative, technology-mediated pedagogies that can adapt to changing circumstances and better cater to diverse student needs.Informed by the complementary theoretical frameworks of Vygotsky's ZPD (Vygotsky, 1978) and Krashen's i+1 theory (Krashen, 1982), the author's approach prioritised the provision of targeted support and scaffolding to foster students' language acquisition within an elective speaking course. Vygotsky's influential theory emphasises the crucial role of social interaction and guidance from more knowledgeable individuals in facilitating learning. Vygotsky (1978) conceptualised the ZPD as the dynamic space between a learner's current independent problem-solving abilities and their potential level of development attained through guidance from more knowledgeable adults or collaboration with more capable peers. This theoretical framework emphasises the pivotal role of social interaction and scaffolding in facilitating an individual's cognitive growth and skill acquisition. The ZPD has been defined as 'the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance, or in

collaboration with more capable peers' (Vygotsky, 1978, p. 86). Drawing on this principle, the author envisioned a digital platform as a means to recreate dynamic learning environments where students could engage in self-directed practices and receive personalised feedback from both instructors and peers.

Krashen's theory of second language acquisition posits that the process is highly analogous to the way children acquire their first language. Krashen (1982) emphasises the pivotal role of meaningful interaction and natural communication in the target language, where the primary focus is on conveying and understanding messages rather than on the grammatical structure of the language. Krashen asserts that error correction and the explicit teaching of linguistic rules are not essential for successful language acquisition. Instead, students learn best when they are able to concentrate on the purpose of communication rather than the form of the language. This theory draws parallels between the language modifications observed in 'foreigner talk' and 'motherese'. Native speakers naturally adapt their speech to be shorter, less complex, and more comprehensible when interacting with non-native speakers or young children learning their first language. These modifications, which often involve the use of distinctive pitch and intonation, are believed to facilitate the language acquisition process by rendering the input more accessible to the learner. Krashen's seminal 'comprehensible input' hypothesis is central to his theory. This hypothesis suggests that language learners acquire new skills by processing and understanding input that is slightly beyond their current level of proficiency, a concept Krashen refers to as 'i+1'. Examples of 'foreigner talk' and 'motherese' are cited by Krashen as evidence of the provision of comprehensible input that is just within the learner's ZPD, thus optimising the conditions for successful language acquisition.

Krashen's i+1 theory further enriched the conceptual foundation of this study by underscoring the importance of comprehensible input. According to this framework, language learners optimally acquire new skills when exposed to linguistic elements that are slightly beyond their current proficiency level (i+1). By strategically integrating Krashen's i+1 theory into the design of the digital platform as an action research initiative, the author aimed to provide students with carefully curated linguistic input that would challenge their current abilities without overwhelming them. The synergistic integration of Vygotsky's and Krashen's theories within the digital platform empowered students to engage in self-directed language practice while simultaneously receiving the necessary scaffolding and comprehensible input from more knowledgeable sources. This multifaceted approach sought to create a dynamic, learner-centred environment that would foster students' language acquisition and progression towards more advanced proficiency levels.

By developing this digital platform, the author aimed to harness the affordances of blended learning, which seamlessly integrates face-to-face instruction with technology-mediated components. This hybrid approach promised to optimise the hybrid learning experience, enabling students to access a wealth of pronunciation tips, multimedia materials, and self-directed practices online while maintaining meaningful in-person interactions with their instructors and classmates offline. Through this action research initiative, the author sought to explore the efficacy of the digital platform in enhancing the teaching and learning outcomes within the elective speaking course. By closely examining the implementation process, analysing student engagement and performance data, and gathering quantitative and qualitative feedback, the author aspired to contribute valuable insights to the ongoing discourse surrounding the strategic integration of digitalisation and blended learning in language

education in higher education. The development of a virtual voice laboratory grounded in the principles of constructivism and second language acquisition theories can significantly enhance students' language learning experience and foster their progression towards greater proficiency, as evidenced through this action research approach.

2. Literature Review

2.1. Action Research in ESL Contexts

The integration of digital tools and blended learning approaches has become increasingly critical for enhancing students' language learning experiences and fostering their progression towards greater proficiency. This importance is underscored by a growing body of action research studies that have explored the efficacy of such practices in diverse EFL contexts. During the COVID-19 pandemic, Chiablaem (2021) explored online tools for English language teaching, with Google Applications for Education (GAFE) emerging as a popular collaborative learning platform. This quantitative study examines Thai university students' positive experiences using GAFE applications like Google Classroom, Meet, Docs, and Forms to enhance English skills in an online course, providing a guideline for integrating these tools into future online teaching. Suryani et al. (2024) investigated the effectiveness of integrating digital technology, specifically the Storyboard That platform, in enhancing English writing proficiency and participant perceptions among nursing students. Employing a quantitative methodology, the research yields significant gains in both knowledge and attitudes, with N-gain scores of 0.53 and 0.52, respectively. The findings underscore the transformative potential of digital tools in augmenting writing skills and fostering a dynamic learning environment, advocating for the seamless integration of technology into English language education. This study serves as a call for educators to embrace the power of digital technology in nurturing a generation of digitally literate and proficient learners. Complementing these scholarly findings, Anis and Khan (2023) explored how multimodal techniques, including visual aids, technology, and creative activities, can promote inclusive practices in the English language classroom. Through a literature review and analysis, the study examines the theoretical foundations and practical implications of using multimodal approaches to enhance inclusivity, engagement, and language acquisition among diverse learners. Collectively, these action research studies underscore the much-needed role that digital tools and blended learning approaches can play in enhancing students' language learning experiences. By providing opportunities for peer feedback, targeted pronunciation practice, interactive communication, self-directed learning, and multimodal expression, these pedagogical approaches have demonstrated their potential to catalyse tangible improvements in students' linguistic proficiency and overall communicative competence. The development of a virtual voice laboratory grounded in these principles can thus serve as a transformative tool that elevates students' language learning and equips them with the skills to thrive in an increasingly digitised and globally interconnected world.

2.2. What Action Research Offers

Action research is well-suited to examining the specific needs, challenges, and dynamics of language teaching contexts within Asian university settings. As Creswell (2012) noted, action research is inherently focused on 'a local problem in a local setting' (p. 576), allowing

educators to deeply engage with the nuanced realities of their own teaching environments. Burns (2010) further emphasised that action research enables language teachers to become 'insider researchers' (p. 2), granting them privileged access to the complex social, cultural, and educational factors that shape the learning experiences of their students. This contextualised approach provides highly relevant data and findings that can directly inform and improve instructional practices tailored to the unique needs of Asian university classrooms (Kemmis et al., 2014).

The cyclical nature of action research, with its iterative processes of planning, action, observation, and reflection, empowers educators to rapidly identify and address the evolving needs of their diverse student populations in Asian universities. Lewin's (1946) influential model of action research highlights the importance of this ongoing adaptive process, enabling teachers to continuously refine their pedagogical approaches in response to emerging challenges and learner feedback. As Nunan and Bailey (2009) asserted, action research allows language teachers to 'respond to the needs and interests of their students' (p. 227), fostering more inclusive and effective instructional practices within the dynamic contexts of Asian higher education. Additionally, the action research process empowers teachers to become active researchers and reflective practitioners, fostering their own professional growth and autonomy within their teaching environments. Zeichner and Noffke (2001) emphasised that engagement in action research helps teachers 'develop a research orientation to their practice' (p. 299), encouraging them to critically analyse their own instructional approaches and experiment with innovative teaching methods. This shift towards a researcher's mindset, as described by Lankshear and Knobel (2004), enables Asian university teachers to become 'self-directed, self-evaluating, and self-renewing' (p. 9) professionals, ultimately enhancing their pedagogical expertise and autonomy.

Action research 'allows teachers to take ownership of the theories and research findings that are presented to them' (Nunan & Bailey, 2009, p. 227), enabling them to critically examine and refine these theoretical frameworks within their own teaching contexts. Hence, action research provides a meaningful bridge between theoretical concepts in language pedagogy and their practical application in real-world Asian university classrooms, enhancing the relevance and impact of scholarly work. This symbiotic relationship between theory and practice, as described by Burns (2010), 'generates knowledge that is relevant to the local setting' (p. 2), ultimately strengthening the connections between language education research and the lived experiences of teachers and learners in Asia. The context-specific nature of action research allows Asian university teachers to develop and test innovative instructional approaches that are tailored to the unique cultural, linguistic, and educational dynamics of their settings. Kemmis et al. (2014) emphasised that action research 'takes place in the here and now' (p. 4), encouraging educators to engage in a process of 'local theorising' (Carr & Kemmis, 1986, p. 165) that is responsive to the specific needs and challenges of their teaching environments. This locally-driven innovation, as described by Nunan and Bailey (2009), can then be shared and adapted by other educators facing similar challenges within the region, fostering a culture of context-sensitive pedagogical development in Asian higher education.

Action research often involves collaborative partnerships between teachers, students, and researchers, facilitating the exchange of insights and the co-construction of knowledge within the Asian academic community. Kemmis et al. (2014) highlighted the importance of this collaborative approach, which 'brings people together to participate in the processes of inquiry

and the improvement of practice' (p. 5). As described by Lankshear and Knobel (2004), language teachers, learners, and researchers in Asia engaging in this shared process of discovery and problem solving can collectively contribute to the development of more effective and inclusive instructional practices. Action research also enables the perspectives of students traditionally underrepresented in Asian universities, such as those from minority linguistic or socioeconomic backgrounds, to be elevated and incorporated into pedagogical improvements. As Creswell (2012) noted, action research 'focuses on giving a voice to marginalised groups' (p. 583), providing a platform for these students to share their unique experiences and shape the direction of instructional development. This emphasis on inclusivity, as described by Nunan and Bailey (2009), ensures that language teaching and learning in Asia is responsive to the diverse needs and perspectives of all learners, fostering more equitable and transformative educational outcomes.

Action research is a vital methodology for enhancing pedagogical practice within Asian university contexts. By examining in depth the specific needs, challenges, and dynamics of their local teaching environments, including the unique circumstances precipitated by the COVID-19 pandemic, educators are able to generate highly contextualised insights that directly inform and improve instructional approaches. Burns (2010) highlighted the pivotal role of action research in improving teaching and learning in diverse educational settings, particularly in the Asia-Pacific region. The value of action research lies in helping teachers develop a deeper understanding of their own practices and the unique challenges within their local teaching environments (Nunan & Bailey, 2009). Action research also enables teachers to investigate the contextual factors that shape their instructional decision-making, especially during times of disruption and uncertainty (Farrell & Jacobs, 2016). Through cyclical processes of planning, action, observation, and critical reflection, action research empowers Asian university teachers to rapidly identify and address the evolving learning needs of their diverse student populations, particularly within hybrid learning modalities. This was demonstrated by Zainal and Yunus (2022), who demonstrated in a systematic review which examined Asian university students' perspectives on online English courses during COVID-19. Key themes from this study included positive and negative views, as well as suggestions to improve online English learning, such as enhancing technological infrastructure and training.

Zeichner and Noffke (2001) emphasized how action research fosters teacher agency, autonomy, and professional development through cycles of planning, action, observation, and reflection, and Bradbury-Huang (2010) highlighted the transformative potential of action research in enabling Asian teachers to become active researchers and reflective practitioners. This reflective praxis fosters teachers' professional autonomy and continuous development as active researchers and scholarly practitioners. Furthermore, action research provides a meaningful bridge between theoretical concepts in language pedagogy and their practical application in real-world Asian university classrooms. It also allows for the development of instructional approaches tailored to the unique linguistic and cultural dynamics of Asian academic contexts (Stenhouse, 1975; Canagarajah, 1999). This enhanced relevance and impact of scholarly work is complemented by the ability of action research to facilitate localised innovation, where instructional approaches are tailored to the unique cultural, linguistic, and educational dynamics of each setting. The collaborative nature of action research also promotes the exchange of insights and the co-construction of knowledge within the Asian academic community (Stringer, 2013). Notably, this approach elevates the perspectives of traditionally

underrepresented students, such as those from minority linguistic or socioeconomic backgrounds, thereby amplifying marginalised voices and informing more inclusive and equitable pedagogies (Phan & Locke, 2016). Having conducted an extensive research, the author recognized the contextual needs and the plethora of benefits brought by conducting action research while building the VVL. The commencement of the action research was driven by the contextual needs to integrate technology in language teaching and learning specified in the following section.

2.3. Technology Integration in Language Learning

The interactive elements that facilitated practical exercises and peer-to-peer interaction in this study were shaped by the author's efforts prior to the COVID-19 pandemic. Nevertheless, even in traditional in-person environments, eliciting active student participation in dialogue-based practice exercises proved challenging. Strategies such as structured pair/group work, task-based learning, incorporating student interests, lowering the affective filter, modeling and scaffolding, and using positive reinforcement were effective in promoting face-to-face interaction. However, the transition to an online learning environment significantly complicated the application of these strategies. While the majority of students provided positive feedback in end-of-term evaluations, several issues arose during the delivery of the speaking course. Before the pandemic, a substantial portion of instructional time was dedicated to facilitating peer-to-peer interaction, allowing the instructor to circulate among groups and provide immediate feedback. In contrast, during virtual instruction, this component was significantly reduced to optimize teaching time, as virtual breakout room discussions often proved ineffective. Consequently, student interaction during class decreased by half, highlighting the challenges of maintaining engagement in a remote context.

Additionally, the operational difficulties of adapting previously successful interactive strategies to an online format exacerbated common errors faced by English as a second language (ESL) students. In breakout rooms, many students defaulted to their first language, complicating instructors' attempts to monitor discussions and ensure students remained on task. Although students were required to keep their cameras on during Zoom classes, it became increasingly difficult for instructors to determine whether they were focused on the lesson or engaging in unrelated activities. This further hindered real-time engagement during group work and made it challenging to provide meaningful feedback. The physical separation inherent in virtual classrooms impeded effective communication and targeted feedback, underscoring the necessity for action research to investigate quality technology integration in language learning.

Facilitating interactive and engaging learning experiences for ESL students in online environments presents significant challenges, thus emphasizing the need for educational action research. As the course leader of an elective speaking course, the author recognized an urgent need for change. Fullan (2007) underscores the critical role of teachers in educational reform, asserting that meaningful change must occur at the teacher level. He highlights the significance of effective professional development, collaborative practices, and a supportive school culture in empowering educators to implement new strategies and improve student outcomes. To initiate this meaningful change, the author decided to conduct action research, aligning with the concept of professional capital outlined by Hargreaves and Fullan (2012). Action research provides a systematic approach to investigate and address the unique challenges arising in

virtual classroom settings, focusing on the efficacy of technology integration in language teaching. This method enables instructors to gain deeper insights into the factors hindering effective interactive and self-directed online learning while developing localized, context-specific solutions. The cyclical nature of action research—comprising planning, action, observation, and reflection—addresses the evolving challenges posed by online learning environments. By bridging theory and practice, action research can foster innovative solutions that consider the pedagogical and technological dynamics of the ESL classroom. Ultimately, this approach aims to enhance the quality of learning experience for students, ensuring that their learning experiences are engaging and effective in both virtual and hybrid contexts. Through this initiative, the author seeks to empower both teachers and students, fostering continuous self-directed and blended learning while producing meaningful educational outcomes.

2.4. Theory Informed Pronunciation Pedagogy in Higher Education

Informed by Krashen's i+1 input hypothesis and Vygotsky's ZPD theory, the author recognized the needs to compensate the reduced interaction between the teacher and the students in the virtual zoom classrooms, and the scaffolding process much needed in peer-to-peer interaction. Before conducting the action research and winning a teaching development grant to build the Virtual Voice Laboratory, the author attempted different means to address the problems arose in the virtual classroom, such as the reduction in teacher-student and peer-to-peer interaction. To compensate for the reduction in physical social interaction, the author redesigned the session structures and assigned individual after-class reading aloud practice for each student in each session. The students were asked to record themselves reading an assigned text aloud and to upload the recording to the university's Moodle learning platform. The instructor then provided written feedback to individual students and posted comments on Padlet.com. The use of Padlet, an additional digital platform to the university's Moodle learning platform, allowed transparency and provided indirect motivation for students to learn how their peers had performed in the reading aloud task. All of the students welcomed this practice of separating recording submission and recording comments. Some of them expressed in the end-of-term course evaluation that they appreciated the transparency of this practice, which allowed the comments to be openly available for all to view. Comments for all students in all sessions taught by the instructor were housed on the same Padlet page.

While the instructional adaptations involving self-directed reading aloud practice and digital feedback platforms were positively received by students, they also introduced several significant challenges. Each instructor's course enrolled between 80 to 100 students, resulting in a weekly requirement to listen to and provide feedback on 160 to 200 individual student recordings. Even if only 70% of students consistently submitted their recordings, this still amounted to 112 to 140 submissions that the instructor needed to address.

Allocating approximately 10 minutes per recording for listening and feedback translated to an additional 18 to 24 hours of work each week dedicated to these non-graded, at-home assignments. After implementing this revised session structure for two semesters during the COVID-19 pandemic, the author recognized that this approach was ultimately unsustainable, both in terms of maintaining instructional quality and managing workload effectively. The significant increase in workload raised concerns regarding the feasibility and scalability of this

practice. Moreover, it became impractical to expect the teaching team to consistently adhere to this model, leading to disparities in the experiences of students enrolled in sessions taught by different instructors. Such inconsistencies highlighted the urgent need for a more efficient and scalable solution that could preserve the benefits of self-directed practice and digital feedback while alleviating the workload concerns for the author and her teaching team.

In response, the author consulted extensive scholarly literature and best practices, proposing the development of a digital platform to facilitate blended and self-directed learning. The challenges faced by the author in managing the extensive feedback workload can be analyzed through the lenses of Vygotsky's Zone of Proximal Development (ZPD) and Krashen's i+1 theory. Vygotsky's ZPD emphasizes the importance of providing support to learners as they navigate tasks that are just beyond their current capabilities, facilitating their growth through guided interaction. In this context, the detailed feedback provided on student recordings represents a critical form of support, enabling learners to progress within their ZPD. However, the unsustainable workload for the instructor limits the effectiveness of this support, potentially hindering students' development if they do not receive timely and consistent feedback.

Similarly, Krashen's i+1 theory posits that effective language acquisition occurs when learners are exposed to input that is slightly beyond their current level of comprehension, allowing them to stretch their understanding without becoming overwhelmed. In the scenario described, while the individualized feedback aims to provide the necessary "i+1" input, the sheer volume of recordings and the subsequent time required for feedback creates a bottleneck in the learning process. As a result, the intended benefits of self-directed practice may be undermined if students do not receive the timely, nuanced feedback necessary for their growth. Hence, the use of the self-directed practice platform needs to be combined with individual consultation sessions to realize the proximal benefits brought by blended learning – with students conducting self-practices outside of classroom and seeking feedback from the course instructors during one-on-one consultation sessions.

By proposing the development of a digital platform to streamline feedback, the author seeks to create a more efficient learning environment that adheres to both Vygotsky's and Krashen's theories. This platform could automate aspects of feedback, ensuring that students receive the support they need within their ZPD while also providing the "i+1" input that fosters language development. Ultimately, this approach aims to balance the demands on instructors and enhance the learning experiences of all students, creating a scalable, effective model for language education that aligns with foundational educational theories.

2.5. Conducting Action Research via a TDG Funded Project

Having identified the pressing needs to enhance teaching and learning efficacies and facing the challenges in teaching and learning in the post-COVID time, the author has undertaken a strategic intervention grounded in the principles of action research (Mertler, 2017; Stringer, 2013). Concomitantly, the researcher has developed a specialised digital platform, the Virtual Voice Laboratory (VVL), to facilitate a more dynamic and personalised learning experience for students. This parallel effort of conducting action research and developing the VVL, supported by a Teaching Development Grant (TDG), has enabled the author to investigate three key research questions. The author seeks to examine the extent to which the VVL platform enhances students' pronunciation skills within the elective course context. Drawing on

Krashen's (1985) i+1 input hypothesis and Vygotsky's (1978) Zone of Proximal Development (ZPD) theory, the VVL has been designed to provide students with comprehensible input slightly beyond their current level of proficiency, while scaffolding their learning within their ZPD. The author also wishes to explore students' perceptions of the effectiveness and usability of the VVL platform in supporting their pronunciation learning. This line of inquiry is crucial in understanding the user experience and potential barriers to adoption, which can inform future refinements and implementation strategies (Hew & Cheung, 2014; Wang et al., 2019). The implementation of the action research also aims to identify the key challenges and considerations in integrating the VVL platform within the broader curriculum and teaching practices. This co-concurrent and holistic perspective is essential in understanding the systemic implications and sustainable pathways for technology integration in language education, particularly in the context of the evolving post-pandemic landscape (Hodges et al., 2020; Rapanta et al., 2020).

2.6. Research Questions

Based on the need to investigate the challenges of facilitating interactive and engaging online learning experiences for ESL students, the following three research questions guided this action research study on the usefulness of a digital platform for blended learning, self-directed practice, and evaluation in a speaking course.

1. To what extent does the implementation of a digital platform for blended learning, self-directed practice, and evaluation enhance the engagement and participation of ESL students in a speaking course?

This research question aims to explore the impact of the digital platform on student engagement and active participation, which are crucial for facilitating interactive and collaborative learning in online and blended learning contexts.

2. How effectively does the digital platform support the development of ESL students' speaking skills, including fluency, pronunciation, and communicative competence, in the speaking course?

This question focuses on investigating the platform's ability to support the targeted skill development of ESL students, which is the primary objective of the speaking course, and how it compares to traditional instructional approaches.

3. What are the perceptions and experiences of ESL students and the instructor regarding the usefulness and challenges of the digital platform in the speaking course's blended learning environment?

This question seeks to gain insight into diverse perspectives, including both students' and the instructor's, to better understand the perceived benefits, limitations, and areas for improvement of the digital platform in the context of the speaking course. By addressing these research questions, action research can provide valuable insights into the effectiveness of the digital platform in enhancing ESL students' engagement, skill development, and the overall learning experience within the blended learning environment of the speaking course. The findings can inform the refinement and optimisation of the digital platform, as well as the broader instructional strategies used in ESL speaking courses.

3. Materials and Methods

3.1. Mechanism of Digital Assistance

Virtual Voice Lab provides a self-directed practice environment for students to engage with the assigned text, 'The Rainbow Passage'. The interface allows students to record their own reading of the passage and compare their waveform to a sample waveform provided. By clicking on the waveform provided (as shown in Figure 1), students can listen to the sample recording. The sample waveform displayed represents the audio recording and the visualisation helps students to find the stressed syllables and the pauses. Students can use the 'Record' function to capture their own reading of the passage, then compare their waveform to the provided model by clicking 'Image of soundwave'. This self-directed practice enables students to self-assess their pronunciation and intonation in relation to the reference recording. The lab aims to facilitate student learning and pronunciation mastery through this interactive, multimodal practice format. Users are instructed to contact technical support if they encounter any issues while using Virtual Voice Lab.

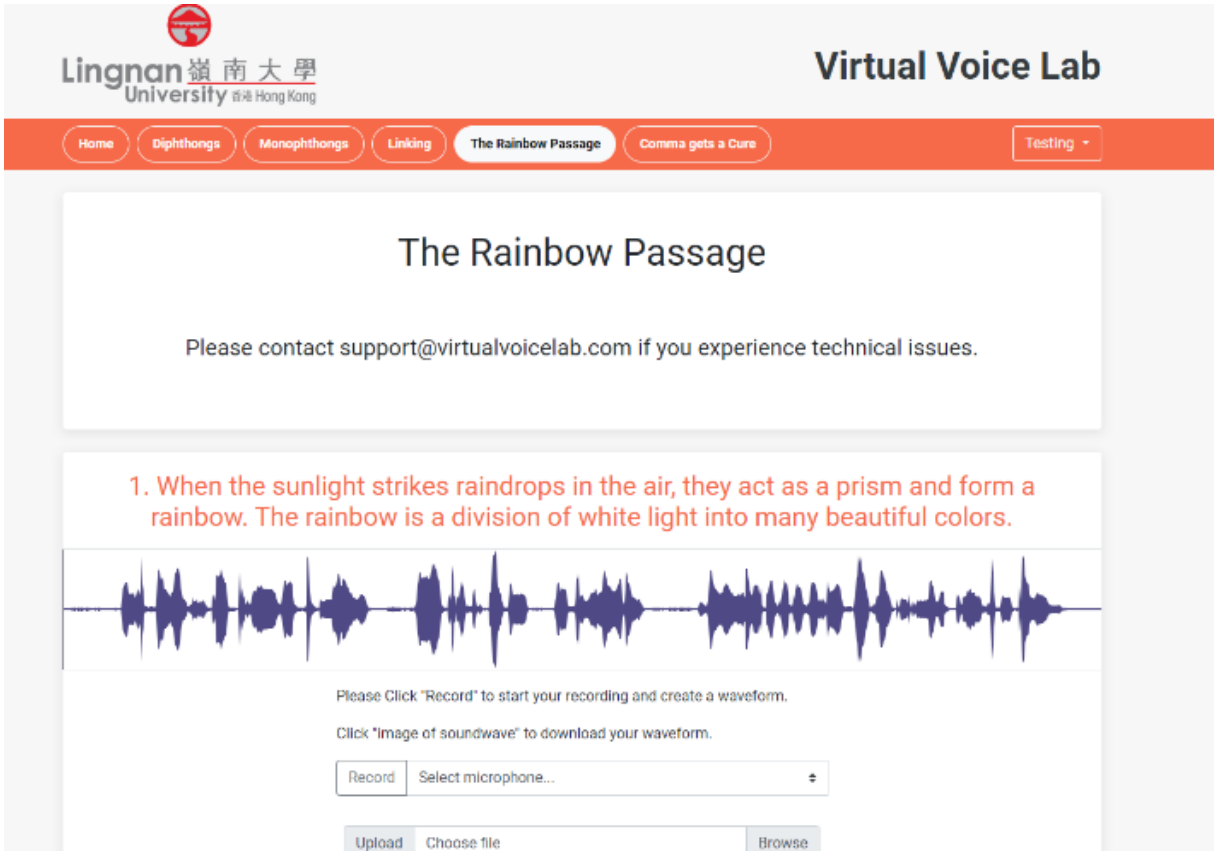

Figure 1. Self-directed Practice Page

Note. The figure shows the self-directed practice page on which students can click the sample waveform to listen to the sample recording. They can then click the 'Record' button to record their reading aloud of the same text.

3.1.1. Analysis of the Students' Waveforms

In the second pilot study, the author invited 79 students enrolled in the four classes to test the virtual voice laboratory, among whom 14 students volunteered to participate in the pilot. Of

these, seven students provided 1-minute video reflections, and nine completed a survey. From the student recordings collected, three specific examples were selected to illustrate how the digital elements of the virtual laboratory facilitated self-directed learning.

The waveform in Figure 2 shows a carefully paced and controlled reading, with clear enunciation of sounds. However, the waveform also indicates an unnatural or stilted delivery, with pauses and pacing that suggest the speaker is not speaking in a natural, conversational manner. The notes provided state that 'The stressed syllables are not aligned, e.g. beautiful. Some ending consonants are missing, e.g. prism, light'. This suggests the recording demonstrates some of the challenges that the user faced in terms of proper stress, rhythm, and articulation when reading the passage aloud. Overall, this recording seems to illustrate the digital lab's ability to provide feedback to users on areas for improvement, such as more natural pacing, stress, and consonant production. The visual waveform allows the user to self-assess their performance and identify specific aspects they need to work on.

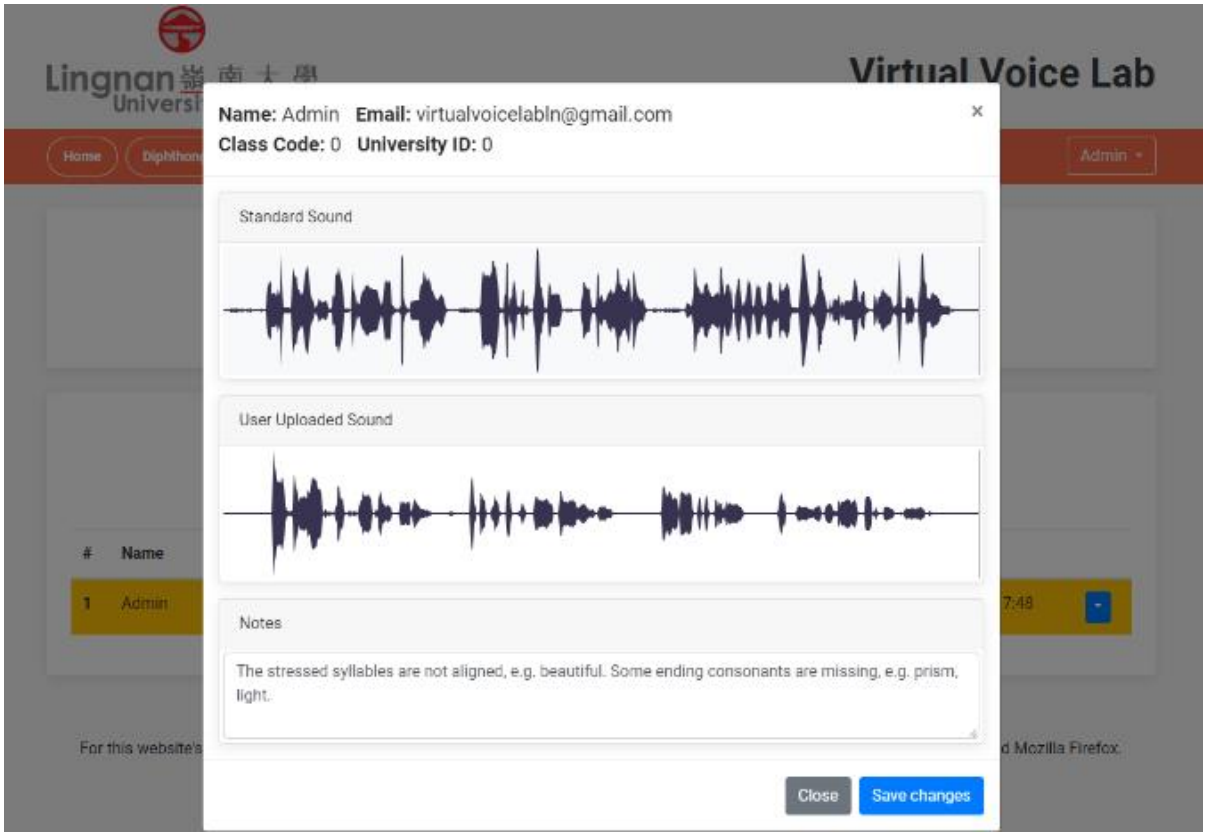

Figure 2. Self-directed Practice Comparison

The waveform for student sample 2 (Figure 3) indicates a much more natural and expressive reading of the passage. The waveform shows clearer delineation between stressed and unstressed syllables, as well as better pacing and rhythm compared to the first recording. The notes provided state that 'The voice chart shows that this reading is overstressed, with too many pauses. But the contrasting waveforms show that there is indeed stressed and unstressed syllables'. This suggests that although the reading may sound a bit overly careful or deliberate at times, the digital lab is effectively capturing the student's improved awareness and control of pronunciation elements such as stress, rhythm, and intonation. The visual feedback allows the student to self-assess their progress and target areas for further refinement. Overall, this second recording demonstrates the value of Virtual Voice Lab in helping students develop more

natural and expressive speech through guided practice and self-evaluation of their performance.

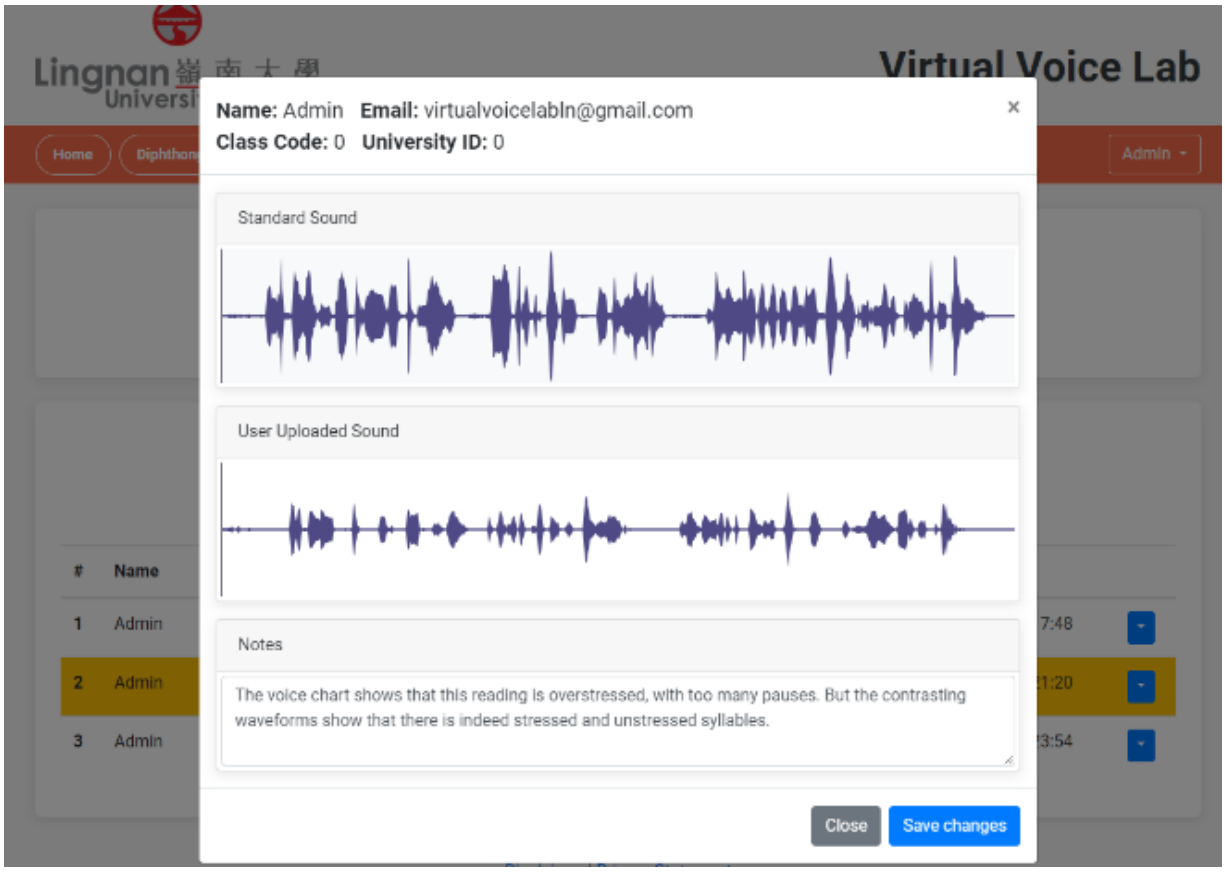

Figure 3. Self-directed Practice Comparison

The waveform for student sample 3 (Figure 4) shows a much more natural and expressive reading compared to the previous two recordings. The pacing and rhythm appear to be well-balanced, with clear delineation between stressed and unstressed syllables. This indicates that although the student has demonstrated good control over pronunciation elements such as stress and articulation, the overall delivery still sounds a bit stilted or unnatural. The slower pacing and deliberate handling of each sound has produced clear enunciation, but at the expense of a more natural, conversational flow. This feedback suggests that Virtual Voice Lab is successfully helping the student identify areas for improvement while also highlighting their progress in mastering the technical aspects of pronunciation. Continuing to use the lab's self-assessment tools will likely allow the student to further refine their reading to achieve a more natural and expressive delivery.

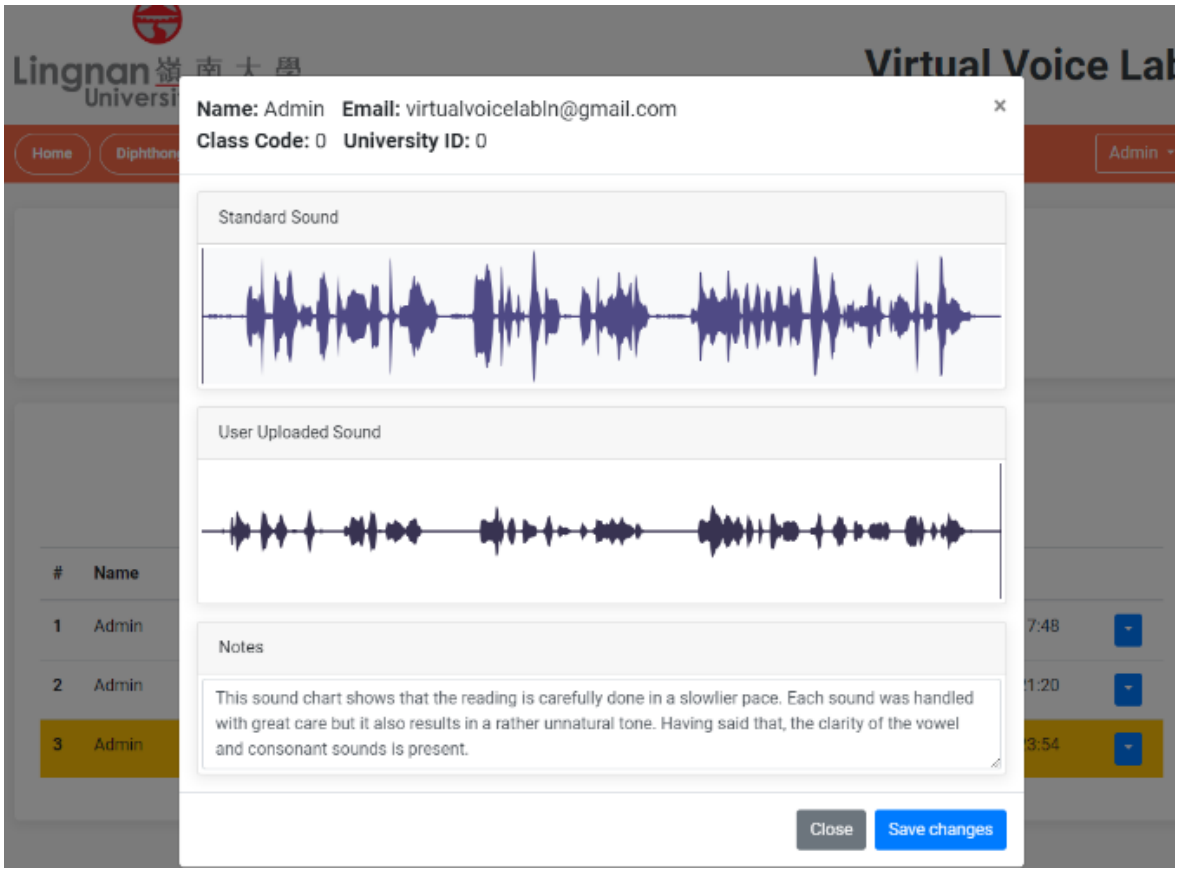

Figure 4. Self-directed Practice Comparison

The selected three student recordings demonstrate various aspects of Virtual Voice Lab's functionality. As shown in the provided waveforms and accompanying notes, the recordings exhibited both strengths and areas for improvement in the students' pronunciation, pacing, and overall expressiveness. The visual feedback provided by the waveform analyses allowed students to self-assess their performance and identify specific areas they could target for further practice and refinement. This self-directed, data-driven approach enabled the students to take an active role in improving their oral reading skills through the virtual lab's interactive features. The pilot study findings suggest that Virtual Voice Lab successfully engaged students in a process of guided, self-paced practice and self-evaluation. By highlighting both the successes and challenges observed in the select student recordings, the author was able to showcase how the digital components of the lab supported students' pronunciation development and autonomy in the learning process.

Two pilot studies were conducted to test the self-directed practice platform 'Virtual Voice Lab' and to collect users' feedback on how to continuously enhance the user experience.

3.2. The First Pilot Study

From December 2022 to April 2023, a total of 640 student-produced audio recordings were collected, assessed, and accompanied by scholarly commentaries. Of these, 10 student recordings were selected for an external consultant to use to test the visualisation coding to complete the backend development of the website. Additionally, the author provided 10 further recordings to serve as model examples for comparison. In November 2023, a pre-pilot test was conducted with 40 students to test the voice visualisation conversion process. Of these, 15

students provided survey feedback on their user experiences and suggestions for improvement. The student input was then compiled and the website content was modified accordingly.

The pre-pilot process undertaken in November 2023 served as a test of the efficacy of the voice visualisation functionality, which is the core digital element enabling students' self-directed learning within Virtual Voice Lab. The insights gained from this initial pilot provided the groundwork necessary to implement this technology-enhanced component successfully. A successful launch of the self-directed learning platform has the potential to address the first challenge that emerged following the collection and assessment of the 640 student recordings by allowing the instructor to efficiently examine the generated sound charts and provide more targeted, individualised feedback to students during consultation sessions. Additional content for blended learning was created from the first quarter of 2024 based on the identified needs of the target student population. The building of Virtual Voice Lab has been on track be launched in the up-coming semester, following the comprehensive groundwork completed during the initial phase of development.

3.3. Data Analysis of the First Pilot Study

Based on the feedback gathered from the 15 students interviewed, several key recommendations were identified to enhance the functionality and usability of the digital self-learning platform within Virtual Voice Lab:

Table 1

	Recommendations given by the student participants on the functionality, usability, and the overall experience in using the VVL
1	Incorporate a feature that allows students to play the demonstration recording and their own recording simultaneously. Synchronised playback would enable learners to more easily compare and evaluate the level of synchronisation between the model pronunciation and their own production, facilitating greater self-assessment and identification of areas for improvement.
2	Include a larger, more prominent button to initiate playback of the demonstration recording. This enhancement would improve the intuitiveness and accessibility of the platform, enabling students to more readily access the model pronunciations during their self-directed practice.
3	Add a table displaying the international phonetic alphabet (IPA) transcriptions alongside the textual content. This feature would provide learners with a supplementary resource to facilitate the accurate identification and production of target sounds, further supporting their self-directed learning and pronunciation practice.
4	Incorporate speech recognition technology to provide automated voice-to-text transcription and analysis of students' recordings. This functionality would allow for the assessment of not only pronunciation accuracy, but also aspects of voice quality and tone, further empowering learners to identify areas for improvement in their oral communication skills.
5	Include a slow-motion playback feature for the demonstration recordings. This enhancement would enable learners to more closely observe and analyse the nuanced articulatory movements and airflow patterns associated with target sounds, thereby strengthening their understanding of proper pronunciation mechanics.

3.4. The Second Pilot Study

Following careful consideration of the student feedback and in-depth consultations with the technical experts responsible for the digital platform's development, the author conducted a

second round of pilot testing involving a larger cohort of participants. In this expanded pilot study, 79 students were invited to engage with Virtual Voice Lab's digital platform over the course of a structured seven-day learning plan. Each day of the plan focused on a specific set of English pronunciation features, including the schwa, stressed syllables, consonant clusters, long and short vowels, diphthongs, and triphthongs. By structuring the pilot study around this focused, sequential approach to pronunciation practice, the project team aimed to provide the student participants with a more comprehensive and immersive learning experience. This systematic method allowed learners to concentrate on mastering individual phonetic elements before progressing to more complex linguistic features, thereby facilitating a deeper understanding and internalisation of the target pronunciation skills.

The decision to implement this structured seven-day plan was informed by the initial round of student feedback, which highlighted the learners' desire for a more guided, scaffolded approach to self-directed pronunciation practice. By addressing this need, the author sought to enhance the overall efficacy and engagement of Virtual Voice Lab, ultimately empowering students to achieve greater improvements in their English articulation and communication abilities. The expanded pilot study, with its emphasis on structured, sequential learning, represents a crucial step in the iterative development process of the digital platform. The insights and data gathered from this second round of user testing will inform further refinements and enhancements to Virtual Voice Lab, ensuring that the final product is optimised to meet the specific needs and learning preferences of the Lingnan University student population. In response to the students' recommendations, an additional seven-day plan was developed. Both seven-day plans will be launched to all students enrolled in the speaking course in the coming academic year and will be implemented during the two weeks during which students meet the instructor for a 10-minute consultation session.

3.5. Data Analysis of the Second Pilot Study

Based on self-reflection scripts from seven volunteers, a thematic analysis was conducted. Four major themes emerged: improvement in pronunciation skills, challenging aspects, recommendations and feedback, and practical aspects.

Table 2

Themes	Description
Improvement in Pronunciation Skills	1. The students generally found the seven-day plan helpful for improving their pronunciation skills, particularly in areas such as linking, stressed syllables, diphthongs, and triphthongs. 2. Several students mentioned experiencing 'obvious improvement' and becoming more confident and effective readers after going through the programme.
Challenging Aspects	1. Certain pronunciation features remained challenging, such as vowels, consonants, and the melody/intonation of sentences. 2. The complexity of linking rules, with exceptions and variations based on accents and speech speed, was identified as a particularly difficult aspect.
Recommendations and Feedback	1. Most students expressed that they would recommend the seven-day plan to their friends and classmates, as they found it useful and beneficial. 2. Suggestions were made to provide sample recordings or answers for comparison to help students better understand the target pronunciation.

Themes	Description
	3. Some students mentioned that the plan would be more suitable for those who have already studied the relevant pronunciation concepts in their coursework.
Practical Aspects	1. The students appreciated the structure of the plan, which allowed them to focus on one pronunciation skill per day and to practise regularly. 2. The ability to record and compare their own voice against the provided recordings was seen as a helpful feature. 3. Some students noted that the system only allowed for one recording per day, which they found limiting.

Overall, the thematic analysis of the student participants' feedback suggested that the seven-day plan and the self-directed practice using Virtual Voice Lab were generally well-received by the students, who recognised its benefits in improving their pronunciation skills, while also highlighting areas for potential improvement and considerations for the target audience.

4. Results

4.1. Survey Results

The findings from the 'Virtual Voice Lab: Seven-Day Plan' survey can be effectively analyzed through the frameworks of Vygotsky's Zone of Proximal Development (ZPD) and Krashen's i+1 input hypothesis. Vygotsky's ZPD emphasizes the importance of providing learners with support that enables them to engage with material that is just beyond their current capabilities. In the context of the seven-day pronunciation improvement plan, the targeted nature of the program likely served to place participants within their ZPD, facilitating their ability to practice pronunciation skills that they were on the cusp of mastering. The positive feedback from participants suggests that the program effectively supported their development, indicating that they received the necessary scaffolding to enhance their pronunciation skills.

Correspondingly, Krashen's i+1 hypothesis posits that effective language acquisition occurs when learners are exposed to input that is slightly above their current proficiency level. The seven-day plan likely offered participants the appropriate "i+1" input, allowing them to practice and refine their pronunciation in a structured way. The participants' generally positive reactions and ratings of the program's effectiveness suggest that the combination of technology-driven tools and focused practice provided them with the necessary input to facilitate improvement.

Overall, the integration of these educational theories underscores the potential of targeted pronunciation pedagogy to enhance language learning in higher education contexts. By aligning the language teaching methods with the principles of ZPD and i+1, the initiative not only addresses the immediate needs of learners but also fosters their ongoing development in a hybrid and effective manner. The survey results illustrate the positive learning experience provided by the 7-day plan and the VVL. In Figure 5, the data clearly indicates that the student participants perceived the 7-day pronunciation improvement plan as effective in enhancing their pronunciation skills. This positive assessment suggests that the structured activities and resources provided throughout the week successfully addressed the participants' needs for improvement in their English pronunciation. The figure presents a range of responses that highlight the participants' experiences, showcasing a consensus among them regarding the plan's effectiveness. The feedback implies that the various elements of the online platform—

such as targeted practice, feedback mechanisms, and the use of technology—collaborated to create an engaging and supportive learning environment. As a result, the participants felt more confident in their pronunciation abilities, which underscores the plan's overall success in meeting its educational objectives. This affirmation not only reflects the participants' satisfaction but also reinforces the potential for such structured methods to facilitate significant improvements in speaking proficiency among learners.

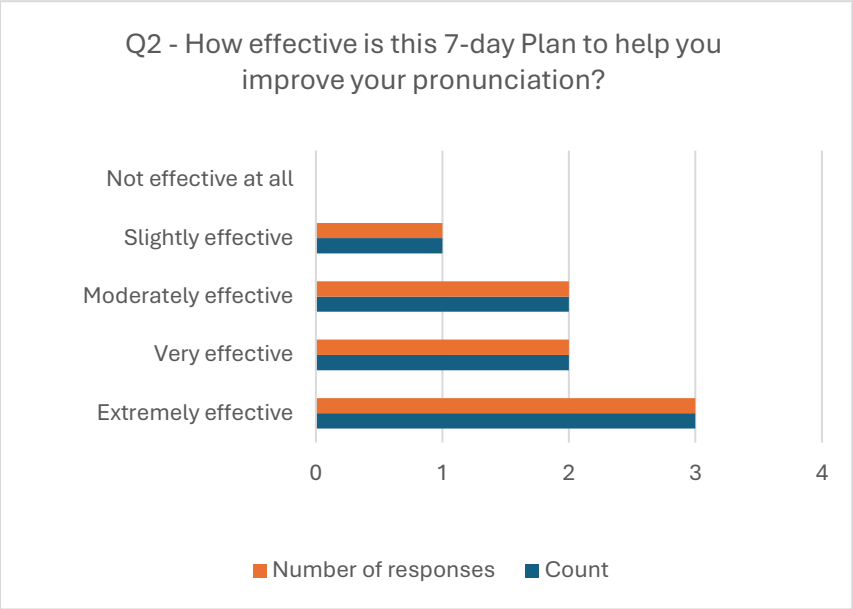

Figure 5. Students' responses of Q2

When responding to Question 3, the student participants highlighted that the most beneficial sessions of the program were Day 3, which focused on "Stressed Syllables and Words," and Day 6, dedicated to "Consonant Clusters." This feedback suggests that these particular topics resonated well with the participants and contributed significantly to their understanding and improvement in pronunciation. As illustrated in Figure 6 below, the emphasis on stressed syllables likely provided students with a clearer grasp of intonation patterns, which are crucial for effective communication. Similarly, the practice on consonant clusters may have equipped learners with the necessary skills to navigate complex sound combinations that often pose challenges in English pronunciation. The positive reception of these days indicates that participants found the content particularly relevant and applicable to their learning needs, reinforcing the importance of targeted instruction in enhancing language acquisition. This data not only reflects the preferences of the students but also serves as valuable insight for future course planning and refinement.

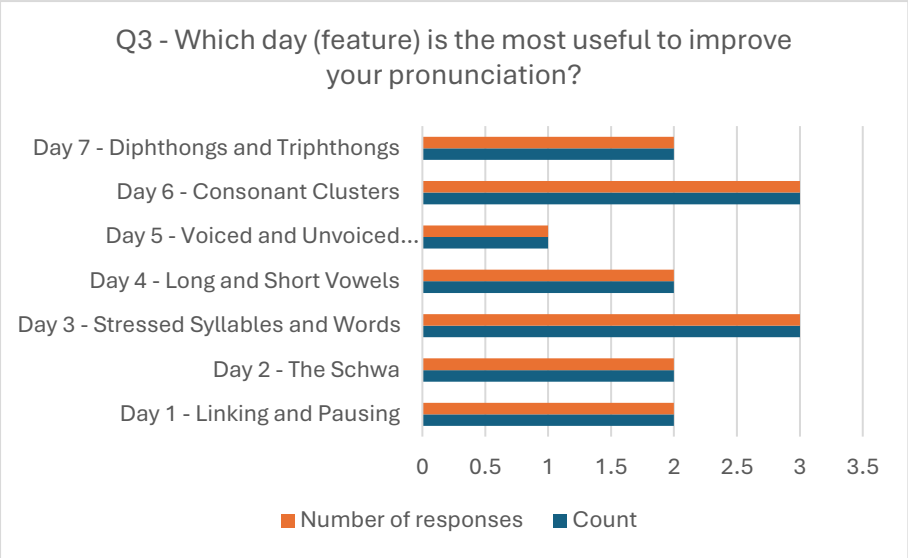

Figure 6. Students' responses of Q3

When asked to rate the helpfulness of the demo recordings, the participants overwhelmingly expressed that they found the recordings to be very helpful in enhancing their understanding and application of pronunciation techniques. Specifically, for the "Clarity" metric, the average rating reached an impressive 4.43 out of a possible 5 stars, based on feedback from 7 respondents. This high score indicates that the participants perceived the clarity of the recordings as a significant asset in their learning process.

In addition, for the "Pace" metric, the average rating was slightly lower but still commendable at 4.00 out of 5 stars, based on 5 responses. This suggests that while participants appreciated the pacing of the demo recordings, there may be room for improvement in this area. The data presented in Figure 7 clearly illustrates these positive perceptions, highlighting the overall effectiveness of the demo recordings in supporting participants' pronunciation practice. Such feedback is invaluable, as it not only underscores the strengths of the demo recordings but also provides insights into specific aspects that could be refined to further enhance the learning experience.

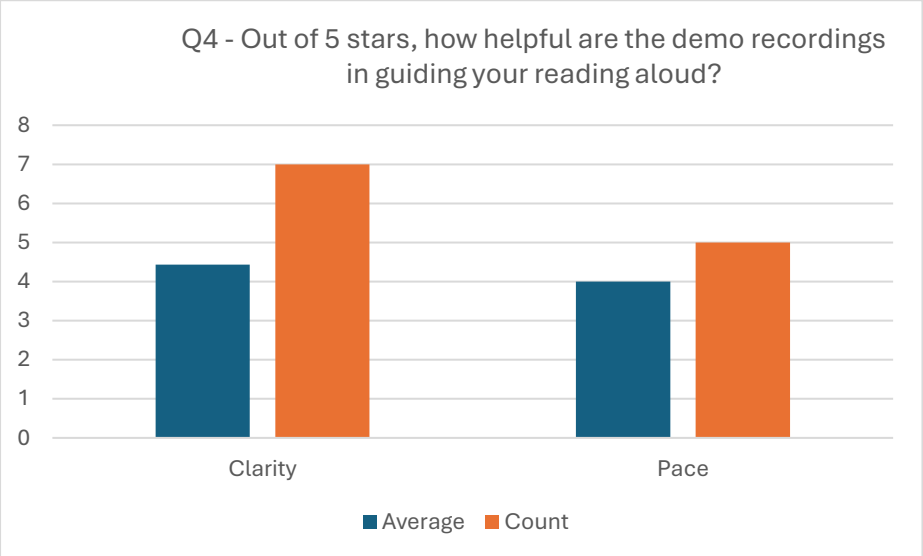

Figure 7. Students' responses of Q4

Aligned with the generally positive reaction noted in Question 2, Figure 8 illustrates that a significant majority of participants expressed a strong likelihood of recommending the seven-day pronunciation improvement plan to their friends. Specifically, the data reveals that most participants indicated they were either extremely likely or somewhat likely to endorse the plan, suggesting a high level of satisfaction with the program's effectiveness. This tendency not only reflects their personal experiences and perceived benefits from the plan but also highlights the potential for positive word-of-mouth promotion, which could encourage more students to engage with the program in the future. Such recommendations may further contribute to the online platform's reputation and effectiveness in enhancing pronunciation skills among language learners.

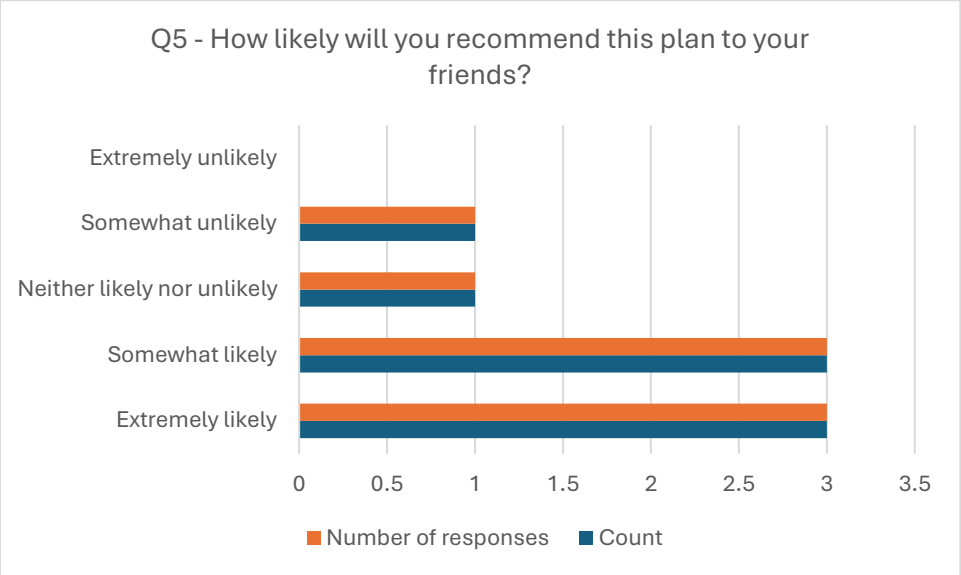

Figure 8. Students' responses of Q5

Question 6 of the survey delved into the ways in which the 7-day pronunciation improvement plan enabled students to engage in independent practice at their convenience. The flexibility of the plan allowed participants to choose their own study times, which is particularly beneficial for accommodating diverse schedules and learning preferences. Additionally, the blended learning approach incorporated visual aids, which proved to be instrumental in helping students identify critical elements such as pauses and stressed sounds in their speech. These visual supports not only enhanced comprehension but also provided a clear framework for learners to follow during their practice sessions.

Most respondents reported that they found the plan easy to navigate, citing the well-organized structure and the accessibility of the materials provided. This ease of use was a significant factor in their ability to practice pronunciation effectively at times that suited them best, as illustrated in Figure 9. The combination of convenience and effective instructional tools contributed to a more engaging and personalized learning experience. The positive feedback indicates that students valued the independence afforded by the plan, which likely increased their motivation to improve their pronunciation skills. Overall, this emphasis on self-directed practice, supported by visual aids, highlights the effectiveness of the 7-day plan in fostering a productive and flexible learning environment.

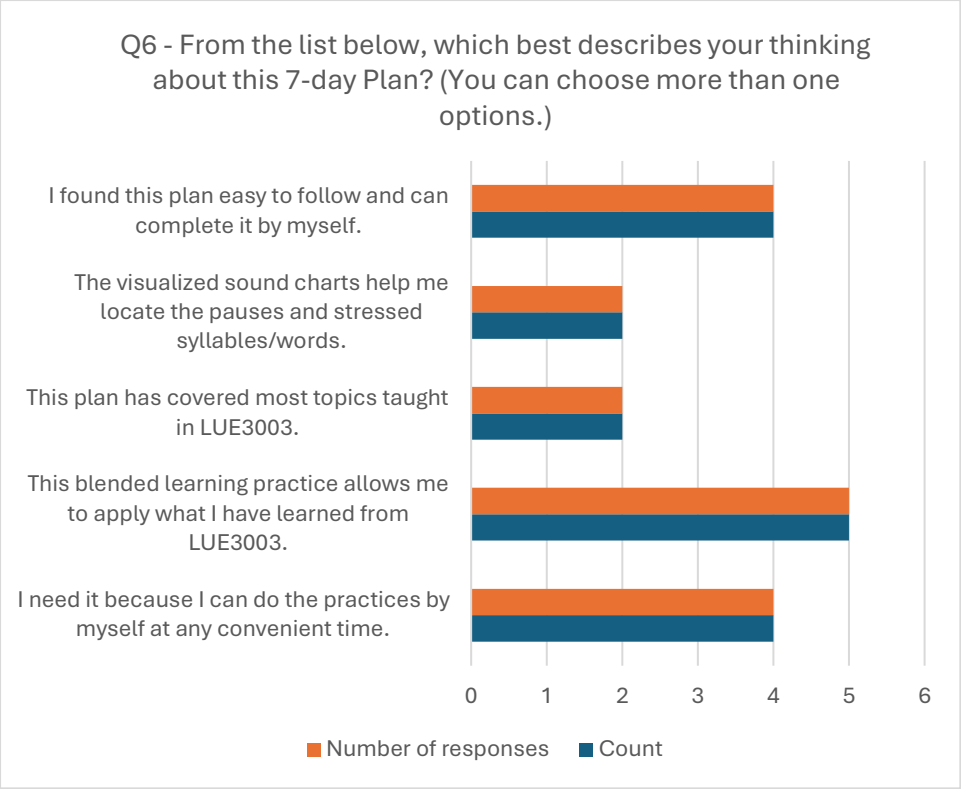

Figure 9. Students' responses of Q6

In Figure 10, the survey results revealed that a significant majority of participants dedicated either 15 to 30 minutes or 30 minutes to 1 hour each day to practicing the activities outlined in the 7-day plan. This finding is particularly noteworthy as it aligns closely with the intended time reduction that was a key aspect of the adapted lesson structure. The structured approach was designed to streamline practice sessions, making them more manageable and accessible for students while still allowing for meaningful engagement with the material.

By committing to these time frames, participants demonstrated their ability to integrate the plan into their daily routines, which suggests that the adjustments made were effective in promoting consistent practice without overwhelming them. This alignment not only indicates that the revised lesson structure was successful in meeting its objectives but also highlights the importance of providing students with realistic and achievable practice durations. Such manageable time commitments likely contributed to the participants' sustained motivation and commitment to improving their pronunciation skills, ultimately enhancing their learning experience throughout the program. Overall, the data underscores the effectiveness of the 7-day plan in facilitating structured, yet flexible, learning opportunities for students.

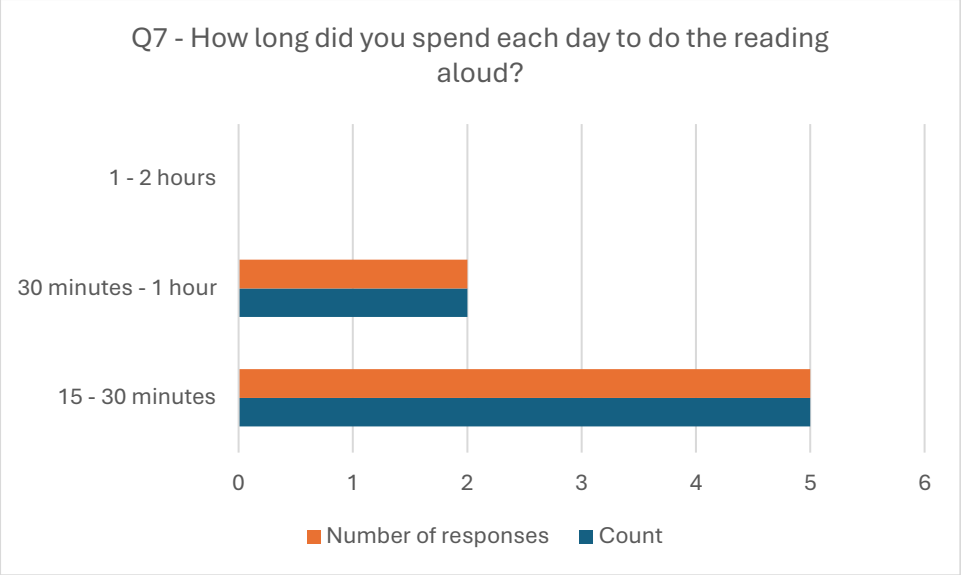

Figure 10. Students' responses of Q7

Some of the participants provided additional comments suggesting that the plan was easy to follow and complete independently by sending the author a self-reflection in video format. In summary, the survey results provide valuable insights into how the Virtual Voice Laboratory (VVL) platform aligns with key language learning theories. The participants' positive ratings of the "Clarity" (4.43 out of 5 stars) and "Pace" (4.00 out of 5 stars) of the demo recordings suggest that the VVL is effectively implementing Krashen's (1985) i+1 input hypothesis. By providing students with comprehensible input that is slightly beyond their current level of proficiency, the VVL appears to be scaffolding their learning within their Zone of Proximal Development, as posited by Vygotsky's (1978) sociocultural theory, with the features of stressed syllables and words in Day 3 and the voiced/unvoiced consonants in Day 6 being the most useful to the student participants. The participants' comments further indicate that the blended learning approach, with its combination of independent practice and in-class support, aligns well with the VVL's design principles. Overall, the survey results demonstrate the VVL's capacity to engage learners and foster their pronunciation skill development through the strategic application of well-established language acquisition theories.

4.1.1. (Dis)alignments of the Reflections and Survey Results

Comparing the survey results and the video reflection feedback revealed several similarities in terms of the students' perspectives on the seven-day pronunciation improvement plan which were summarised in Table 3 below:

Table 3

Category	Similarities
Improvement in Pronunciation Skills	Both sources indicate that the students generally found the seven-day plan helpful in improving their pronunciation, particularly in areas like linking, stressed syllables, diphthongs, and triphthongs.
Challenging Aspects	Both sources indicate that certain pronunciation features, such as vowels, consonants, and the melody/intonation of sentences, remained challenging for the students.
Recommendations and Feedback	Both sources suggest that most students would recommend the seven-day plan to their friends and classmates, as they found it useful and beneficial.

The video reflection feedback provided more nuanced insights compared to the survey data. It indicated stronger degrees of improvement, highlighted the complexity of linking rules, suggested supplementary materials, and commented on the plan's suitability for different proficiency levels and practical features - aspects that were not as prominent in the broader survey findings. This detailed feedback complemented the overall understanding gained from the survey responses as showed in the Table 4.

Table 4

Category	Differences
Degree of Improvement	The video reflection feedback indicates that several students experienced 'obvious improvement' and became more confident and effective readers, whereas the survey results do not provide such a strong emphasis on the degree of improvement.
Complexity of Linking Rules	The video reflection feedback specifically highlights the complexity of linking rules, with exceptions and variations based on accents and speech speed, as a particularly difficult aspect, which is not as prominent in the survey results.
Recommendations for Supplementary Materials	The video reflection feedback suggests providing sample recordings or answers for comparison to help students better understand the target pronunciation, which is not mentioned in the survey results.
Suitability for Different Proficiency Levels	The video reflection feedback suggests that the seven-day plan might be more suitable for students who have already studied the relevant pronunciation concepts in their coursework, while the survey results do not explicitly mention this.
Practical Aspects	The video reflection feedback emphasizes the students' appreciation for the structure of the plan and the ability to record and compare their own voice, while the survey results do not provide such detailed feedback on the practical aspects.

In summary, the survey and the video feedback provide a complementary understanding of the students' experiences of the seven-day pronunciation improvement plan, highlighting both similarities and nuanced differences in their perspectives and recommendations. These similarities and differences also suggest that the seven-day plan was generally well-received, with students finding it helpful in improving their pronunciation skills. However, the video reflection feedback provides more nuanced and detailed insights into the specific challenges faced by the students and their recommendations for further improvements to the plan. Based on the differences between the survey results and the video reflection, the author identified the limitations of the self-directed practice of the VVL and produced a possible solution as follows.

4.2. Limitations of the Self-directed Practice

A key limitation of the digital platform is the potential for delayed and limited instructor feedback due to the platform's scalability challenges. With potentially large class sizes of up to 100 students per semester, an instructor would only be able to provide detailed feedback on a select number of student recordings. This could result in many students receiving delayed or infrequent feedback, diminishing the immediacy and relevance of the guidance. Additionally, the asynchronous nature of the platform, where students record and submit their work for later review, means there is a lack of real-time interaction between the instructor and students. This could prevent students from receiving timely interactive feedback and guidance that could help them make immediate adjustments and improvements. While the digital platform provides

valuable self-assessment tools and opportunities for independent practice, these scalability issues may ultimately limit the quality and timeliness of instructor engagement, potentially undermining the platform's effectiveness in fully supporting students' self-directed learning.

Another notable limitation of the project stemmed from the restricted funding available, which constrained the development of the digital platform. Because of financial constraints, the platform's functionality was limited to self-diagnosis by students and manual assisted diagnosis by instructors. A significant portion of the available funds was allocated to backend development and staffing costs to support the external technical consultant who provided expert guidance on the platform's development. Given the limited financial resources, it was not feasible to develop a fully automated self-diagnostic system at this stage of the project.

4.3. Solution: Accompany Self-directed Learning with Blended Learning

The author believes that the ability to read aloud with clarity, proper pronunciation, and expressive delivery can significantly enhance one's articulation. To address this need and help students make better use of the self-directed platform to compensate for the lack of self-diagnostic function, the author created a seven-day reading aloud plan that presents a structured self-guided programme aimed at improving English pronunciation through systematic daily practice. In the context of using the Virtual Voice Lab (VVL) to facilitate self-directed learning, balancing feedback with assessment is crucial for enhancing student engagement and promoting effective learning outcomes. Nicol and Macfarlane-Dick (2006) emphasize that feedback should be intricately linked to assessment, allowing students to understand how their performance aligns with established learning objectives. Through the VVL, students can record their speaking practice and with a digital record of their own voice chart, students can receive targeted feedback when meeting the course instructors that not only addresses their current performance but also guides them toward achieving desired standards. This feedback mechanism working under the blended learning environment enables students to identify specific areas for improvement while fostering self-assessment skills, as they can reflect on their recordings and compare them against the provided benchmarks. Moreover, the VVL promotes opportunities for students to close the gap between their existing capabilities and their goals by practicing at a pace that works best for them. After having practiced online, outside classroom, the course instructors can offer detailed, constructive feedback tailored to individual needs. This approach encourages a dialogue around feedback, as students can engage with instructors and peers to discuss insights derived from their recordings, thereby deepening their understanding of effective pronunciation and specific stress and intonation patterns. By integrating this self-practice and targeted feedback into subsequent assessments, the VVL reinforces the principle that feedback given in individual consultation sessions is not merely evaluative but serves as a vital tool for ongoing learning. As students apply the feedback to improve their future recordings, they develop a clearer sense of progression and mastery, which enhances their motivation and self-efficacy. Furthermore, the instant comparison of the two voice charts within the VVL cultivates a learning environment where students feel empowered to take ownership of their learning journey. Ultimately, by effectively balancing feedback with self practices through the use of the Virtual Voice Lab, the course instructors can foster a more hybrid and productive learning experience, ensuring that students not only meet learning objectives but also develop the skills necessary for lifelong language learning on their own. This integration of self-directed learning and blended learning thus becomes a cornerstone

of a comprehensive approach to teaching speaking, promoting both immediate improvement and long-term growth.

The core text of this instructional plan is 'The Rainbow Passage', which is a widely recognised text that has been used extensively in the field of speech and language studies. This passage was originally developed by George Shipp and published in the 1930s as a tool for evaluating and enhancing speech production (Fairbanks, 1960; IDEA, 2021).2021). It has gained widespread recognition and use within the field of speech and language therapy, as well as in language learning and assessment contexts. The text is particularly well-suited for this purpose due to its rich linguistic features, including a variety of vowel sounds, consonant blends, and diphthongs, as well as variations in stress and intonation patterns. The author chose the text based on a remarkable project undertaken by the International Dialects of English Archive (IDEA). The IDEA team has collected recordings of 'The Rainbow Passage' being read aloud by individuals from diverse linguistic and cultural backgrounds, representing a wide range of English accents and varieties from around the world. The purpose of this project is to showcase the inherent beauty and diversity of English pronunciation, challenging the notion of a single 'standard' or 'correct' way of speaking the language. By assembling a collection of 'The Rainbow Passage' readings from speakers of different nationalities, regional dialects, and sociolinguistic backgrounds, the IDEA project aims to celebrate the richness and dynamism of the English language as it is spoken by people from all corners of the globe.

By using the same text and with detailed instructions, learners are guided to focus on specific phonetic and suprasegmental features of English over seven consecutive days. These include the schwa sound, which is a critical but often challenging aspect of English vowel quality, as well as the identification and production of stressed and unstressed syllables within words. Additionally, the plan emphasises the accurate articulation of long and short vowels, as well as the distinction between voiced and unvoiced consonants. Learners also explore the complexities of consonant clusters and the production of diphthongs and triphthongs, which are characteristic of English phonology. Equally important, the programme addresses the skills of linking words and strategically pausing during oral reading, which are crucial elements of prosodic fluency. The plan outlines a step-by-step approach, encouraging learners to record their readings on the VirtualVoiceLab.com platform and provide feedback on their progress. This multifaceted approach, which combines structured practice, self-evaluation, and optional external resources, is intended to facilitate meaningful improvement in the learners' oral reading abilities. By following this seven-day plan, English language learners and speakers can develop a deeper understanding of the nuances of English pronunciation, empowering them to communicate more effectively in academic, professional, and social contexts. The systematic nature of the programme, coupled with its emphasis on targeted practice and self-reflection, makes it a valuable tool for individuals seeking to enhance their overall linguistic proficiency.

Despite the financial limitations that precluded the development of a fully automated self-diagnostic system, a comprehensive seven-day instructional plan was designed to systematically guide users in practising and improving their English pronunciation. This structured approach, with specific guidelines and a focus on different aspects of pronunciation each day, was intended to facilitate progress in reading aloud skills over the course of one week. By leveraging the detailed seven-day plan, the project sought to compensate for the inability to implement a fully automated self-diagnostic system.

4.4. Launching Virtual Voice Lab

In the next phase of the action research, the author will invite all 220 students enrolled in the elective course to engage in self-directed practice during the scheduled consultation weeks. The course is structured over a 14-week semester, with two designated consultation weeks during which students will have the opportunity to meet with the instructor individually for 10-minute sessions to receive personalised assistance. To facilitate self-learning and promote a blended learning approach, all students will have access to a range of supporting materials, including instructional videos on diphthongs, triphthongs, linking, and individual sounds, and interviews will be held with students discussing their experiences with challenging English phonemes. These resources are designed to complement the structured seven-day plan and to empower learners to take an active role in their pronunciation development.

To ensure the completion of the self-directed learning component, it will be mandatory for students to upload their practice recordings to the designated platform. The instructor will then listen to one recording per student and provide brief feedback on their progress. As an incentive, students will be awarded 5% of the attendance and participation mark for successfully completing the seven-day plan using Virtual Voice Lab. This multifaceted approach combining individualised instructor support, access to multimedia learning resources, and a structured self-directed practice component aims to foster a learning environment that is both engaging and effective. By encouraging students to take ownership of their pronunciation development and providing them with the necessary tools and guidance, the author hopes to facilitate meaningful improvements in their oral communication skills.

5. Conclusion

5.1. Driving Digitalization into the Future

This action research study has important implications for both theory and practice in the field of ESL instruction and educational technology. From a theoretical perspective, the findings build upon Vygotsky's zone of proximal development (Vygotsky, 1978) and Krashen's input hypothesis (Krashen, 1985) by demonstrating the value of using digital tools and structured instructional plans to support students' language development within their ZPD. The study also extends the existing research on the integration of technology to enhance teaching and learning in higher education (Al-Maroof et al., 2020). By actively engaging students and incorporating their feedback, the study illustrates the importance of grounding instructional design in the experiences and needs of language learners, in line with principles of learner-centered pedagogy (Coman et al. 2020; Zainal & Yunus, 2022).

In terms of practical implications, this study provides a model for how ESL instructors can leverage digital platforms and tailored instructional plans to address the challenges of shortened attention spans among students (Hossen & Uddin, 2023). The detailed insights from the student feedback highlight specific areas of language development that remain challenging, as well as recommendations for supplementary materials and scaffolding to support learners. These findings can inform the ongoing refinement and improvement of the instructional plan and the use of the Virtual Voice Lab platform.

Future research should further explore the efficacy of the seven-day instructional plan and the

Virtual Voice Lab platform in supporting ESL learners' language development across different proficiency levels and contexts. Additional studies could investigate the long-term impact of such interventions on learners' language skills and confidence, as well as their transferability to real-world communication. Comparative studies examining the relative benefits of technology-mediated and traditional language instruction would also contribute valuable insights. Ultimately, this action research study demonstrates the value of an iterative, learner-centered approach to designing and refining instructional practices in the field of ESL education and educational technology.

Funding

This project was funded by UGC's Teaching Development Grant 2022-24 (Project code: 102710).

References

Al-Maroof, R. S., Salloum, S. A., Hassanien, A. E., & Shaalan, K. (2020). Fear from covid-19 and technology adoption: The impact of google meet during coronavirus pandemic. *Interactive Learning Environments*, *31*(3), 1293–1308. https://doi.org/10.1080/10494820.2020.1830121

Anis, M., & Khan, R. (2023). Integrating multimodal approaches in English language teaching for inclusive education: A pedagogical exploration. *Universal Journal of Educational Research, 2*(3), 241-257. https://doi.org/10.5281/zenodo.8365506

Bowen, J. A. (2012). *Teaching naked: How moving technology out of your college classroom will improve student learning*. Jossey-Bass, a Wiley imprint.

Bradbury-Huang, H. (2010). What is good action research? *Action Research*, *8*(1), 93–109. https://doi.org/10.1177/1476750310362435

Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners*. Routledge. https://doi.org/10.4324/9780203863466

Canagarajah, A. S. (1999). *Resisting linguistic imperialism in English teaching*. Oxford University Press.

Carr, W., & Kemmis, S. (1986). *Becoming critical: Education, knowledge and action research.* Falmer Press.

Chiablaem, P. (2021). Enhancing English Communication Skills of Thai University Students through Google Apps for Education (GAFE) in a Digital Era during Covid-19 Pandemic. *Shanlax International Journal of Education*, *9*(3), 91-98. https://doi.org/10.34293/education.v9i3.3921

Coman, C., Țîru, L. G., Meseșan-Schmitz, L., Stanciu, C., & Bularca, M. C. (2020). Online teaching and learning in higher education during the coronavirus pandemic: Students' perspective. Sustainability, *12*(24), 10367. https://doi.org/10.3390/su122410367

Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.). Pearson.

Fairbanks, G. (1960). *Voice and articulation drillbook*. (pp.127). New York, NY: Harper and Row.

Farrell, T. S. C., & Jacobs, G. M. (2016, January 31). Practicing what we preach: Teacher reflection groups on cooperative learning. *TESL-EJ*. https://eric.ed.gov/?id=EJ1092795

Fullan, M. (2007). *The new meaning of educational change* (4th ed.). Routledge.

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, *7*(2), 95-105. https://doi.org/10.1016/j.iheduc.2004.02.001

Garrison, D. R., & Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines.* Jossey-Bass. https://doi.org/10.1002/9781118269558

Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *The handbook of blended learning: Global perspectives, local designs* (pp. 3-21). Pfeiffer.

Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning*, *18*(4), 311-326. DOI: https://doi.org/10.1080/09588220500335455

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Routledge.

Hew, K. F., & Cheung, W. S. (2014). *Using blended learning: Evidence-based practices*. Springer. https://doi.org/10.1007/978-981-287-089-6

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*, 27. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Hossen, M. K., & Uddin, M. S. (2023). Attention monitoring of students during online classes using XGBoost classifier. *Computers and Education: Artificial Intelligence*, *5*, 100191. https://doi.org/10.1016/j.caeai.2023.100191

IDEA. (2021). *The rainbow passage: Idea: International dialects of English archive*. IDEA: International Dialects of English Archive. https://www.dialectsarchive.com/the-rainbow-passage

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer. https://doi.org/10.1007/978-981-4560-67-2

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Longman.

Krashen, S. D. (1985) *The Input Hypothesis: Issues and Implications.* London: Longman.

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, *20*(3), 271-289. DOI: https://doi.org/10.1017/S0958344008000335

Lankshear, C., & Knobel, M. (2004). *A handbook for teacher research: From design to implementation.* Open University Press.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, *2*(4), 34-46. https://doi.org/10.1111/j.1540-4560.1946.tb02295.x

Mertler, C. A. (2017). *Action research: Improving schools and empowering educators* (5th ed.). SAGE Publications. https://doi.org/10.4135/9781483396484

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218. https://doi.org/10.1080/03075070600572090

Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. Heinle Cengage Learning.

Oblinger, D. G., & Oblinger, J. L. (Eds.). (2005). *Educating the net generation.* EDUCAUSE.

Phan, L. H., & Locke, T. (2016). 'I'm not the same after my research': A teacher's personal experience of the transformative power of practitioner research. *Changing English*, 23(4), 364-378.

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2020). Online university teaching during and after the Covid-19 crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education*, *2*(3), 923-945. https://doi.org/10.1007/s42438-020-00155-y

Reimers, F. M., & Schleicher, A. (2020). *A framework to guide an education response to the COVID-19 Pandemic of 2020*. OECD. Retrieved from: https://www.oecd.org/education/A-framework-to-guide-an-education-response-to-the-Covid-19-Pandemic-of-2020.pdf

Salmon, G. (2013). *E-tivities: The key to active online learning* (2nd ed.). Routledge. https://doi.org/10.4324/9780203074640

Sharma, P., & Barrett, B. (2007). *Blended learning: Using technology in and beyond the language classroom.* Macmillan.

Stenhouse, L. (1975). *An introduction to curriculum research and development.* Heinemann.

Stringer, E. T. (2013). *Action research (4th ed.)*. Sage Publications.

Suryani, N. Y., Rizal, S., Rohani, T., & Ratnaningsih, H. (2024). Improving learners' English writing skills through digital technology and Project-Based Learning. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, *8*(1), 21–34. https://doi.org/10.22437/jiituj.v8i1.32506

Thorne, S. L., & Payne, J. S. (2005). Evolutionary trajectories, Internet-mediated expression, and language education. *CALICO Journal*, *22*(3), 371-397. https://doi.org/10.1558/cj.v22i3.371-397

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wang, Y., Huang, X., Schunn, C. D., Zou, Y., & Ai, W. (2019). *Redesigning flipped classrooms: A learning model and its effects on student perceptions - higher education*. SpringerLink. https://doi.org/10.1007/s10734-019-00366-8

Zainal, N., & Yunus, Md. M. (2022). Asian University Students' perspectives on Online English courses during COVID-19: A systematic review. *International Journal of Evaluation and Research in Education (IJERE)*, *11*(2), 888. https://doi.org/10.11591/ijere.v11i2.22420

Zeichner, K., & Noffke, S. (2001). Practitioner Research. In V. Richardson (Ed.), *Handbook of Research on Teaching* (pp. 298-332, 4th ed.). Washington DC: American Educational Research Association.