INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 14, Issue 4, (2025) E-ISSN: 2226-6348

Seeing the Voice: Spectrogram Based Visual Feedback in Teaching Mongolian Long Song in Higher Education

Zhai Haoyan, Jinky Jane C. Simeon

Academy of Arts and Creative Technology (ASTiF), Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia

Email: jinkysoo@ums.edu.my

Corresponding Author Email: zhaihaoyan12345@126.com

DOI Link: http://dx.doi.org/10.6007/IJARPED/v14-i4/27038

Published Online: 05 December 2025

Abstract

This study investigates how spectrogram based visual feedback can support the teaching and learning of Mongolian Long Song (Urtiin duu) in higher music education. While Long Song has been recognised as intangible cultural heritage, its core vocal technique, *Nogula*, remains difficult for many university students especially non Mongolian speakers to perceive and control through listening alone. Building on recent research that uses digital audio technology to enhance vocal instruction, this study explores the use of spectrogram visualisation as a pedagogical tool in a Long Song course at an arts university in Inner Mongolia. A qualitative case study design was adopted, combining classroom observations, pre and post lesson audio recordings, spectrogram analysis, student reflective journals and semi structured interviews. During a series of technique lessons, students received visual feedback on their own singing through spectrograms generated in Praat and related software. Data were analysed thematically, with spectrogram patterns serving as complementary evidence to students' and teacher's reflections. The findings suggest that spectrogram based visualisation can help students better understand the line of breath, the internal wave structure of *Nogula* and differences between tense and relaxed vocal production. Students reported increased awareness of their own vocal habits and greater confidence in self correction. At the same time, the study identifies limitations and cautions regarding over reliance on visual representations. Implications are discussed for technology enhanced heritage music education and for integrating acoustic visualisation into progressive vocal pedagogy.

Keywords: Mongolian Long Song, Spectrogram, Visual Feedback, Technology Enhanced Learning, Higher Music Education, Intangible Cultural Heritage

Introduction

In recent years, technology enhanced learning has become a central topic in progressive education, including in the field of music. Digital audio workstations, mobile applications and visualisation tools are increasingly used to support instrumental and vocal instruction, providing students with instant feedback and new ways of engaging with sound (Laurillard,

2012). In higher music education, such tools are commonly integrated into courses that focus on popular and Western classical repertoires, where studio based production, detailed acoustic analysis and digitally mediated performance have already become part of the curriculum (Ban & Bin Md Noor, 2024).

By contrast, much less attention has been paid to how similar technologies might assist the teaching of traditional or heritage vocal genres whose transmission has historically depended on oral practice, apprenticeship and community life. For these genres, the central pedagogical challenge is not only technical accuracy, but also the safeguarding of cultural meanings and local value systems within formal institutions. This raises a broader question for social science and music education research, namely how digital tools can be used in ways that support rather than displace existing modes of learning and identity formation in heritage music contexts.

Mongolian Long Song (Urtiin duu) provides a pertinent case for examining this question. Recognised by UNESCO as Intangible Cultural Heritage of Humanity, Long Song is renowned for its extended melodic lines, wide tessitura and intricate ornamentation (UNESCO, 2003). In the Inner Mongolia Autonomous Region of China, the genre has moved from pastoral settings into urban conservatoires and universities and has become both a subject of formal study and a symbol of Mongolian cultural identity. Its core ornamental technique, Nogula, involves delicate wave like modulations produced through subtle interactions of breath, laryngeal movement and resonance. For many university students, especially those whose schooling has been conducted primarily in Mandarin and who may not speak Mongolian fluently, the fine grained internal motion of Nogula is difficult to perceive and control through listening and imitation alone.

These technical challenges are embedded in a wider institutional context. Long Song teachers working in higher education must negotiate expectations of scientific vocal training and measurable learning outcomes, while also preserving the genre's ecological and cultural meanings (Pilcher, 2018; Wang et al., 2023). Some teachers express concern that, in the absence of appropriate explanatory tools, administrators and colleagues may misinterpret Long Song technique as rough or unhealthy when judged against Western bel canto norms. Others worry that efforts to standardise Long Song for examinations may flatten its internal diversity and detach it from everyday Mongolian life (Smith & Akagawa, 2009; Titon, 2009). The question of how to demonstrate the systematic nature and pedagogical value of Long Song technique, without reducing it to a simplified or decontextualised form, therefore has both practical and theoretical implications.

Spectrograms and other forms of acoustic visualisation offer one potential way to address these concerns. By converting sound into visible patterns of frequency and intensity over time, spectrograms can reveal aspects of vocal production that are difficult to describe verbally, such as the continuity of breath, the shape of vibrato and the timing of ornamentation. Research in speech therapy, pronunciation training and vocal pedagogy suggests that visual feedback can help learners notice differences in resonance and articulation that they might not easily hear (Bliss, 2018; Kartushina et al., 2015; Lã & Fiuza, 2022). At the same time, work on intangible cultural heritage warns that technological interventions should respect community values and avoid reducing complex living practices

to technical data (Smith & Akagawa, 2009; Titon, 2009). This tension indicates a gap in current scholarship regarding how acoustic visualisation can be integrated into heritage vocal pedagogy in a culturally responsive way.

Responding to this gap, the present study examines a case of Mongolian Long Song teaching in which spectrograms were integrated into technique lessons at an arts university in Inner Mongolia. A Long Song bearer teacher collaborated with the first author to experiment with spectrogram based visual feedback as a way of making breath lines and Nogula more tangible for students, while retaining oral explanation, imagery and community grounded metaphors as the core of instruction. The study seeks to understand how spectrogram visualisation is used in practice, how it shapes students' understanding of Long Song technique and how participants perceive its benefits and limitations within the broader ecology of Long Song transmission.

This study addresses the following research questions:
RQ1: How is spectrogram based visual feedback integrated into the teaching of Mongolian Long Song in a higher music education setting?
RQ2: In what ways does spectrogram visualisation influence students' understanding and control of key Long Song techniques, particularly *Nogula* and the breath line?
RQ3: How do students and the teacher perceive the benefits and limitations of using spectrograms as a learning tool in Long Song classes?

By exploring these questions, the article seeks to contribute to ongoing discussions about technology enhanced learning, culturally responsive music education and the sustainable transmission of intangible cultural heritage in formal institutions.

The study is novel in two main respects. First, it extends research on technology enhanced learning and spectrogram based feedback to a heritage vocal tradition that has been under represented in empirical studies, thereby showing how acoustic visualisation can be adapted to an oral, community rooted pedagogy. Second, it contributes to social science discussions of culturally responsive and sustainable music education by demonstrating how digital tools can support learner agency, intercultural access and institutional recognition of Mongolian Long Song, while still foregrounding the ecological and cultural meanings that give the genre its social significance.

Literature Review

Technology Enhanced Learning and Visual Feedback in Music Education

Technology enhanced learning has been widely promoted as a means of supporting progressive, learner centred education by personalising instruction and providing richer feedback (Laurillard, 2012). In music, digital tools have been employed for practice tracking, audio visual modelling, online collaboration and automated assessment. Studies of technology enhanced music education report that students appreciate the immediacy, interactivity and multimodality of such tools, which can increase motivation and allow them to practise more flexibly outside class (Ban & Bin Md Noor, 2024).

Within vocal pedagogy, digital audio technology has been used to visualise aspects of vocal production and to provide objective feedback on pitch accuracy, timing and dynamics. For

example, Ban and Bin Md Noor (2024) describe a case study of a modern popular singing course in which digital audio workstations and plug ins were used to show students their pitch contours and spectral balance, helping them to refine intonation and stylistic nuances. They argue that integrating such tools into vocal teaching can empower learners to take more responsibility for their own progress, provided that teachers mediate the technology critically. Although much of this work has focused on Western popular or classical singing, its underlying pedagogical logic using technology to make sound processes more visible and manipulable may also be relevant to traditional vocal genres. However, when working with intangible cultural heritage musics, scholars and practitioners caution that technology should be used in ways that respect community values and avoid reducing complex practices to decontextualised data (Smith & Akagawa, 2009; Titon, 2009).

Spectrograms and Acoustic Visualisation in Vocal and Language Learning

Spectrograms have a long history in phonetics and speech science, where they are used to analyse the acoustic properties of speech and singing (Boersma & Weenink, 2025). In educational settings, spectrograms have been employed to support pronunciation training, speech therapy and, more recently, vocal pedagogy. Bliss (2018) reviews the use of computer assisted visual articulation feedback in second language pronunciation instruction and concludes that spectrographic displays can help learners notice subtle features of speech, such as voice onset time and vowel quality, especially when combined with explicit instruction.

Kartushina et al. (2015) found that phonetic production training with visual feedback improved both the perception and production of foreign speech sounds, suggesting that visual representations can strengthen the link between auditory targets and articulatory control. In vocal pedagogy, Lã and Fiuza (2022) report on a pilot study where real time acoustic analysis and visual feedback were used to support singing training. They argue that visualisation can serve as a complementary tool for teachers, enabling them to explain abstract concepts like resonance and breath support with reference to observable patterns.
In the context of self directed learning, Baagbah and Ganapathy (2025) show that learners using Praat to visualise lexical stress in English developed greater awareness of their prosodic patterns and reported increased autonomy. Together, these studies suggest that spectrograms and related tools can act as “visual scaffolds” that help learners connect what they feel in the body with what they hear and see.

However, researchers also note potential drawbacks. Some learners may find spectrographic displays intimidating or overly technical, and there is a risk that students focus on producing visually “perfect” patterns at the expense of naturalness or expressivity (Bliss, 2018; Lã & Fiuza, 2022). The effectiveness of visual feedback appears to depend on careful pedagogical mediation and on integrating technology within broader frameworks of culturally responsive teaching (Gay, 2010; Ladson Billings, 1995).

Heritage Music, Mongolian Long Song and Higher Education

Since the adoption of the Convention for the Safeguarding of the Intangible Cultural Heritage, UNESCO and its member states have recognised the importance of safeguarding living cultural practices, including music, through community based transmission and formal education (UNESCO, 2003). Scholarship on intangible heritage has highlighted tensions between

preservation and innovation, between community ownership and institutional control and between safeguarding practices and the commodification of culture (Smith & Akagawa, 2009; Soini & Birkeland, 2014).

Titon (2009) proposes an ecological view of music and sustainability, in which musical practices are understood as part of wider sound ecologies involving relationships among people, places and other beings. From this perspective, sustaining a musical tradition involves sustaining the social and environmental conditions that allow it to flourish, not merely preserving fixed repertoires or performance styles.

Mongolian Long Song exemplifies many of these issues. Historically, Long Song has been tied to pastoral life, horse culture and seasonal rituals. Pilcher (2018), in an ethnographic study of Urtiin duu in Alshaa, Inner Mongolia, notes that Long Song singing is closely linked to experiences of riding horses and moving across the steppe; older singers often remark that one cannot truly sing Long Song without knowing how to ride. Contemporary research on Long Song includes work on lyrical classification (Namjil, 2022) and on speech signal processing applied to long tune folk songs (Wang et al., 2023), reflecting growing interest in its linguistic and acoustic features.

Within higher education, incorporating Long Song into curricula offers opportunities for cultural sustainability and identity affirmation, particularly for Mongolian students who may have been educated primarily in Mandarin (Gay, 2010). At the same time, the move into conservatoires and universities introduces new pressures to codify technique, standardise assessment and adapt to classroom time structures. As Soini and Birkeland (2014) argue, cultural sustainability requires balancing continuity with change, and respecting community meanings while engaging with new institutional and technological contexts.

Despite this growing body of work, there is still limited research on how digital technologies, such as spectrograms, can be used in pedagogically and culturally sensitive ways to support the transmission of heritage vocal traditions like Long Song. The present study seeks to address this gap by examining a concrete case of spectrogram based visual feedback in Long Song teaching.

Methodology

Research design

This study adopted a qualitative case study design with elements of technology enhanced action research. The first author, a Long Song practitioner and doctoral candidate, collaborated with a Long Song bearer teacher to integrate spectrogram based visual feedback into a series of technique lessons and to document how this intervention shaped teaching and learning processes. The action research component involved iterative cycles of planning, implementing, observing and reflecting on the use of spectrograms, with adjustments made based on students' responses and the teacher's observations.

Context and Participants

The case was situated in a Long Song technique course offered to undergraduate voice majors at a regional arts university in the Inner Mongolia Autonomous Region. The course was part of the institution's ethnic music programme and met once a week for a two hour session over

a four week block. The class was taught by a master Long Song singer and recognised cultural heritage bearer, hereafter referred to by the pseudonym “Aya”, who visited the university as a guest instructor.

A total of 14 students (aged 18–23) participated in the study. Eight identified as Mongolian and six as Han Chinese. All were enrolled in the university’s music performance programme, with primary specialisations in either Western classical voice or Chinese national vocal music. None had previously received systematic Long Song training, although some Mongolian students had heard Long Song in family or community contexts.
Participation in the research component was voluntary. Students were informed that their learning activities would proceed as planned regardless of consent, and that only those who agreed to participate would have their data used in the study.

Spectrogram Intervention and Teaching Sequence
The spectrogram intervention took place during four consecutive lessons focused on basic Long Song technique. The main software used was Praat (Boersma & Weenink, 2025), complemented by a simple spectrogram plug in in a digital audio workstation. The researcher set up a laptop with an external microphone and a projector so that students could see spectrograms in real time or shortly after singing.

In the first lesson, Aya introduced the concept of natural voice and guided students through speaking based exercises, such as reading short sentences and sustaining vowels on a single pitch. She emphasised the importance of finding a comfortable, speech like tone that did not strain the throat. After students sang a short Long Song phrase individually, the researcher recorded each performance and generated a basic spectrogram. Aya then projected selected spectrograms and pointed out the overall breath line and areas where the sound broke or weakened.

In the second and third lessons, the focus shifted to *Nogula*. Aya used the “hei hei” exercise: students produced a series of short, light “hei” syllables, gradually connecting them into longer wave like motions. Spectrograms were used to compare students’ attempts with Aya’s demonstration, highlighting differences between irregular, noisy patterns and more regular internal waves. Students were also encouraged to look at their own spectrograms during individual practice, experimenting with different breath support and laryngeal relaxation.
In the fourth lesson, Aya and the students revisited the initial Long Song phrase and recorded new performances. Spectrograms of pre and post intervention recordings were displayed side by side to facilitate reflection on changes in breath continuity and *Nogula* structure.

Data Collection
Data were Collected Through Multiple Sources
Classroom observations and field notes. The first author observed each lesson, taking detailed notes on Aya’s explanations, gestures, use of spectrograms and interactions with students.
Audio recordings and spectrograms. Individual student performances of a short Long Song phrase were recorded before and after the intervention. Additional shorter clips were recorded during practice to illustrate particular issues. Spectrograms were generated for these recordings and stored for analysis.

Student reflective journals. After each lesson, students wrote short reflections (200–300 words) responding to prompts about what they had learned, how they felt about using spectrograms and what they found challenging.
Semi structured interviews. At the end of the teaching block, Aya and six volunteer students (three Mongolian and three Han) participated in semi structured interviews lasting 30–45 minutes. Interviews explored their experiences with spectrograms, perceptions of changes in singing and views on the relationship between technology and tradition.

Data Analysis
Data analysis proceeded in two complementary stages. First, audio recordings and corresponding spectrograms were examined to identify broad patterns in students' vocal production, such as the continuity of the breath line, the internal wave structure of *Nogula* and the presence of tension related noise. The aim was not to produce detailed acoustic measurements but to use spectrograms as visual evidence that could be triangulated with qualitative data.

Second, interview transcripts, reflective journals and field notes were analysed thematically. Following an inductive deductive approach, the researcher initially coded segments of text related to students' and teacher's experiences with spectrograms, perceived changes in technique, emotional responses and cultural meanings. Codes were then grouped into themes that addressed the three research questions. Throughout the analysis, preliminary interpretations were shared with Aya to check for resonance with her perspectives and to avoid misrepresentation.

Ethical Considerations and Positionality
Ethical approval for the study was obtained from the relevant university committee. Participants were informed about the purpose of the research, the voluntary nature of participation and the measures taken to protect confidentiality. Pseudonyms are used for students in all reporting, and audio files are stored securely.
The first author is herself a Long Song singer, an officially recognised inheritor of Long Song in China and a young lecturer in musicology. This positionality enabled close access to the teaching process and to the embodied aspects of vocal learning, but it also required reflexivity regarding assumptions and potential biases. Regular dialogue with Aya and critical discussions with colleagues helped to foreground students' voices and to treat analysis as a shared interpretive process.

Findings
Four themes illuminate how spectrogram based visual feedback shaped Long Song teaching and learning in this case: (1) making breath lines visible; (2) seeing the internal waves of *Nogula*; (3) supporting self correction and confidence; and (4) negotiating the limits of visualisation.

Making Breath Lines Visible
A first theme concerns how spectrograms helped students visualise the flow of breath across the long phrases characteristic of Long Song. Before the intervention, several students reported that they understood the idea of "one breath for one line" conceptually, but could

not judge whether they were actually sustaining the sound evenly. As one student, Lin (Han, soprano), explained in her first reflection:

> *"Teacher Aya always says we should 'ride one breath to the end of the hill', but I only know if I run out of air or not. I cannot see whether my sound is really like one line or many small pieces."*

When Aya projected the spectrogram of Lin's initial recording, the image showed a clear drop in intensity halfway through the phrase, with the main horizontal band breaking into fragmented patches. Aya traced the discontinuity on the screen and asked Lin to recall how she had felt at that moment. Lin replied that she had slightly panicked and tightened her throat to compensate for lack of air.

Over the next lessons, Lin and her peers practised deeper inhalation and more economical use of breath. In the final session, they compared their pre and post intervention spectrograms. Lin's second recording showed a more continuous horizontal band, with smoother intensity. She commented:

> *"When I saw the two pictures together, I realised that the 'straight road' Teacher Aya talks about is not just a metaphor. The spectrogram really shows whether my breath is continuous. It gives me confidence that I am moving in the right direction."*

Similar patterns were observed for other students. Mongolian student Nara noted that, although she was familiar with Long Song from childhood, she had not previously paid attention to how evenly her breath flowed within a single phrase. Seeing her own spectrogram led her to adjust her posture and to avoid squeezing at the end of lines.

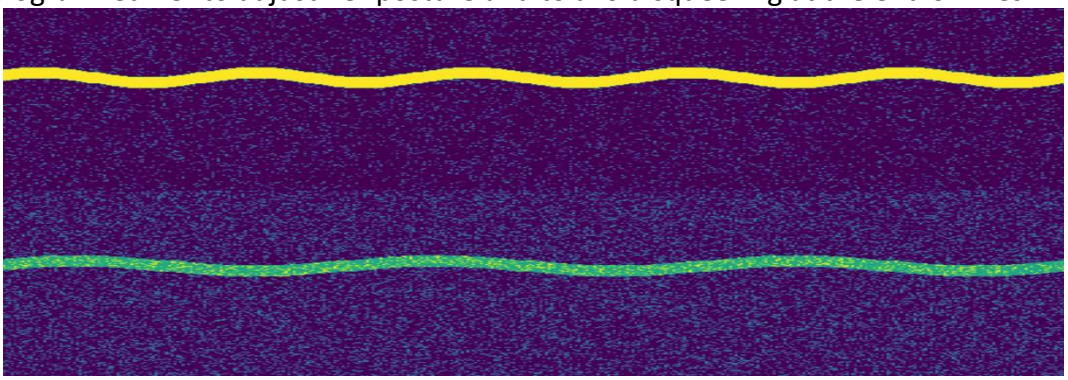

Figure 1. Spectrograms of a student's performance of the same Mongolian Long Song phrase before (top) and after (bottom) the spectrogram based intervention, showing increased continuity of the breath line.

Seeing the Internal Waves of Nogula

The second theme relates to students' understanding of *Nogula*. Many beginners initially tried to imitate the external appearance of *Nogula* by shaking the jaw or forcing the larynx. These attempts produced a harsh, irregular sound and caused discomfort. In the spectrogram, such production appeared as jagged, unstable patterns with scattered noise across frequencies.

Aya contrasted these images with her own demonstration, which showed a stable fundamental frequency with regular, fine grained undulations. The difference was striking even to students with no background in acoustics. Han student Xiaoyu reflected:

> *"Before, I thought that as long as the sound was shaking, it was Nogula. When I saw my spectrogram next to Teacher Aya's, I realised that my pattern was messy and full of noise, while hers was like smooth waves. It helped me understand why my throat felt tight."*

Through repeated practice of the "hei hei" exercise and visual feedback, students began to focus more on letting the breath create gentle internal waves rather than mechanically shaking external muscles. Mongolian student Batu described the shift:

> *"Teacher Aya asked me to imagine that inside one long line there are many small waves, but the sea itself is calm. When I watched the spectrogram, I could see these small waves appearing when I relaxed my throat. It became a game to see how regular I could make them."*

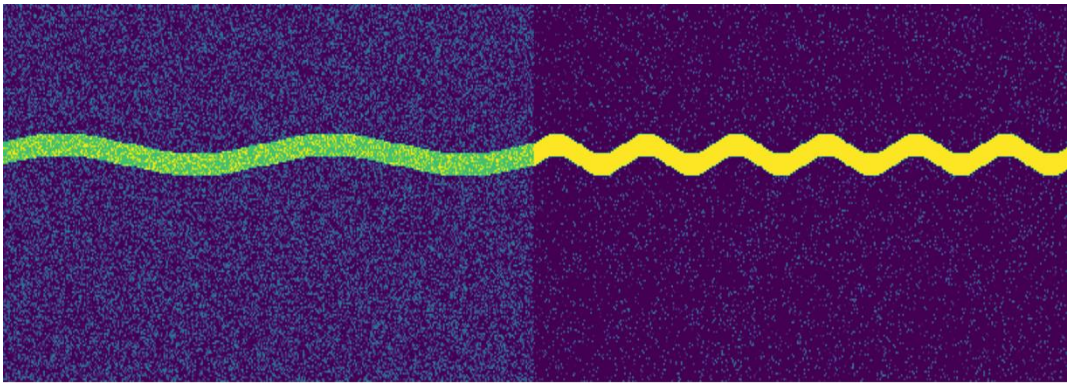

Figure 2. Spectrogram comparison of Nogula: (left) a novice student's initial attempt with an irregular and noisy pattern; (right) the teacher Aya's demonstration with a stable pitch band and regular internal waves.

The visualisation thus served as a bridge between Aya's metaphors (sea, waves, riding breath) and students' bodily sensations, making abstract ideas more concrete without reducing them to rigid formulas.

Supporting Self Correction and Confidence

A third theme concerns the role of spectrograms in supporting self correction and learner autonomy. Once students became more familiar with the basic features of spectrograms horizontal axis as time, vertical axis as frequency, colour or darkness as intensity they started to use the tool independently during practice sessions.

In between formal lessons, several students came to the classroom early to record themselves and inspect the resulting spectrograms. They experimented with different breath strategies and laryngeal positions, observing how the visual patterns changed. For example, Xiaoyu described in her journal:

> *"I recorded the same phrase three times: first with my old 'popular music' voice, then with a very relaxed speech like voice and finally with more breath support. On the spectrogram, I could see that the first one had many random dark spots, the second was too thin and the third was more even. It made me*

feel that I could discover things by myself, not only wait for the teacher's comments."

Students reported that this process increased their sense of control. Rather than fearing that they might accidentally damage their voices, they saw spectrograms as a kind of safety monitor. Nara noted:

"When my throat felt tired, I checked the spectrogram and saw that there was a lot of noise in the high frequencies. After I relaxed my tongue and jaw, the noise became less. It reassured me that I was singing in a healthier way."

Aya also observed positive changes in students' confidence during performance. She felt that those who had engaged actively with spectrograms dared to sustain longer phrases and to explore higher notes, because they had developed a clearer internal model of how their voices behaved.

Negotiating the Limits of Visualisation

Despite these benefits, the final theme highlights limitations and cautions identified by both Aya and the students. Some students initially found spectrograms intimidating. Han student Ming wrote:

"The first time I saw the spectrogram, it looked like a medical test from a hospital. I felt nervous, as if the machine would tell me that my voice was 'bad'."

Aya shared this concern. She worried that over emphasis on visual patterns could lead students to pursue visually "beautiful" spectrograms rather than listening deeply to the timbre and emotional content of Long Song. In one interview, she remarked:

"Long Song is not only about lines and waves. It is about the smell of horses, the colours of the mountains, the feeling of riding. The spectrogram cannot show these things. If students only look at the screen, they may forget to look at the grassland in their hearts."

To address this, Aya deliberately limited the frequency and duration of spectrogram use within each lesson, ensuring that most of the time was still devoted to singing, listening and discussing imagery. She also framed spectrograms as "one more mirror" rather than as an authority, reminding students that the final judge of good singing was how it felt in the body and how it touched listeners.

Another limitation concerned access and sustainability. The intervention relied on a laptop, microphone and projector set up that may not be available in all classrooms, especially in rural or community based teaching contexts. Some students expressed a wish for simpler tools, such as mobile applications, that could offer basic visual feedback without specialised equipment.

These reflections suggest that while spectrograms can be powerful pedagogical aids, their use must be carefully balanced with the values and practical realities of heritage music transmission.

Discussion

The case study sheds light on how spectrogram based visual feedback can be integrated into the teaching of a heritage vocal tradition in higher education and what this implies for technology enhanced music education more broadly.

First, the findings support previous research on spectrograms and acoustic visualisation as effective tools for making invisible aspects of sound tangible for learners (Bliss, 2018; Kartushina et al., 2015; Lã & Fiuza, 2022). In the context of Long Song, spectrograms helped students to see the continuity (or discontinuity) of their breath lines and the internal wave structure of *Nogula*. This aligns with the notion of visual feedback as a scaffold that can bridge the gap between auditory targets and bodily control (Baagbah & Ganapathy, 2025). For students who lacked prior experience with Long Song or with Mongolian language, the visual representation provided an additional channel of understanding that complemented, rather than replaced, listening based learning.

Second, the study illustrates how technology can support culturally responsive music education when used in dialogue with community knowledge (Gay, 2010; Ladson Billings, 1995). Aya's teaching did not treat spectrograms as neutral scientific tools but interpreted them through metaphors rooted in pastoral life such as roads, waves and hills thus connecting digital images to cultural imagery. At the same time, spectrograms offered a way to communicate with institutional colleagues unfamiliar with Long Song, demonstrating that the genre's vocal production is systematic and healthy, even if it does not conform to Western bel canto norms. This may help counter deficit views of indigenous or heritage musics within higher education.

Third, the case highlights the importance of balancing technology with tradition in efforts toward cultural sustainability. From an ecological perspective, sustaining Long Song involves maintaining relationships among singers, landscapes, animals and communities, not merely preserving acoustic patterns (Titon, 2009; Soini & Birkeland, 2014). Aya's concern that students might focus too much on "beautiful" spectrograms points to the risk of reducing Long Song to technical parameters detached from its lived meanings. The intervention suggests that technologies like spectrograms are most helpful when they are used sparingly, in ways that deepen rather than displace embodied, imaginative and relational aspects of learning.

Finally, the study has implications for progressive education and development. Spectrogram based visual feedback appeared to enhance students' sense of agency and self regulation, qualities valued in contemporary frameworks of learner centred education. By enabling students to monitor their own vocal production and to engage in exploratory practice, the intervention supported the development of reflective, independent musicianship. At the same time, it broadened access to Long Song for non Mongolian students, who could use visual cues to enter a culturally specific sonic world. This speaks to the potential of technology enhanced heritage music education to contribute to inclusive and diverse learning environments in higher education.

Conclusion

This article has examined a case of Mongolian Long Song pedagogy in which spectrogram based visual feedback was integrated into higher music education. Through a qualitative case study design that combined classroom observations, audio recordings, spectrogram analysis, student reflections and interviews, the study explored how spectrograms were used, how they influenced students' understanding of Long Song technique and how participants perceived their value and limitations.

The findings suggest that spectrogram based visualisation can play a constructive role in Long Song teaching by making breath lines and *Nogula* more visible, supporting self correction and strengthening students' confidence. When mediated by a culturally grounded teacher, spectrograms can serve as a bridge between traditional oral metaphors and contemporary digital literacies, contributing to culturally responsive and progressive music education.

At the same time, the study underscores that such tools must be used judiciously. Over reliance on visual patterns risks narrowing the focus of learning to technical surfaces and neglecting the ecological, emotional and spiritual dimensions of Long Song. Teachers and institutions should therefore integrate spectrograms within pedagogies that remain centred on listening, imagination and relationships with community and landscape.

Future research could extend this work by involving larger and more diverse student cohorts, comparing different types of acoustic visualisation tools or examining similar interventions in other heritage vocal traditions, such as throat singing or Korean *gagok*. Collaborative projects among Long Song bearers, music technologists and educators may also explore the design of accessible applications that support practice while respecting community values.

In an era when intangible heritage practices are vulnerable to both disappearance and superficial appropriation, experiments like the one documented here suggest that carefully used technologies can help young singers not only to hear and feel Long Song, but also to see their own voices as part of a living tradition.

References

Baagbah, S. Y., & Ganapathy, M. (2025). Examining the use of PRAAT to improve English lexical stress production: Insights from self directed learning within the framework of self determination theory. *Language Teaching Research Quarterly, 49*, 170–193.

Ban, J., & Bin Md Noor, A. R. (2024). Empowering music education: Leveraging digital audio technology in the pursuit of vocal mastery – A case study of the modern popular singing course. *International Journal of Academic Research in Progressive Education and Development, 13*(4), 423–433.

Bliss, H. (2018). Computer assisted visual articulation feedback in L2 pronunciation instruction: A review. *Journal of Second Language Pronunciation, 4*(2), 145–171.

Boersma, P., & Weenink, D. (2025). *Praat: Doing phonetics by computer* (Version 6.4.47) [Computer program]. Praat.org.

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice* (2nd ed.). Teachers College Press.

Kartushina, N., Hervais Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of

foreign speech sounds. *The Journal of the Acoustical Society of America, 138*(2), 817–832.

Lã, F. M. B., & Fiuza, S. M. (2022). Real time visual feedback in singing pedagogy: A pilot study on acoustic analysis in vocal training. *Applied Sciences, 12*(19), 9966.

Ladson Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491.

Laurillard, D. (2012). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.

Namjil, B. (2022). Long song type classification based on lyrics. *Journal of Multimedia Information System, 9*(2), 113–120.

Pilcher, O. W. (2018). "You can't sing urtyn duu if you don't know how to ride a horse": Urtyn duu in Alshaa, Inner Mongolia. *Asian Music, 49*(1), 4–33.

Soini, K., & Birkeland, I. (2014). Exploring the scientific discourse on cultural sustainability. *Geoforum, 51*, 213–223.

Smith, L., & Akagawa, N. (Eds.). (2009). *Intangible heritage*. Routledge.

Titon, J. T. (2009). Music and sustainability: An ecological viewpoint. *The World of Music, 51*(1), 119–137.

UNESCO. (2003). *Convention for the safeguarding of the intangible cultural heritage*. United Nations Educational, Scientific and Cultural Organization.

Wang, K., Li, B., & Xue, Y. (2023). Speech signal processing and analysis based on Mongolian long tune folk songs. *Current Research in Agricultural and Life Sciences, 1*(1), 15–23.