
Audiobooks as Mediators of Engagement and Learning in Online English Education

Audiolivros como Mediadores do Empenho e da Aprendizagem no Ensino Online de Inglês

Susana M. Oliveira¹

Abstract

The article presents an empirical study on the use of audiobooks in online English language learning within a distance higher education context. Developed as part of a broader exploration of sound-based texts, the empirical inquiry reflects on how literary voice can sustain engagement, deepen interpretation, and humanise the experience of asynchronous learning. It also considers how the human voice — singular, authentic, and emotionally resonant — can renew empathy and attention in digital spaces increasingly mediated by technology and artificial intelligence. Drawing on data from 528 adult learners across three consecutive English courses, the research combines participation records, performance indicators, and anonymised qualitative reflections. The learning sequences, based on short stories by O. Henry, Chimamanda Adichie, and Salman Rushdie, interwove multimodal activities involving listening, pronunciation, and reflective writing. Across levels, participation remained stable (71–72%), with mean performance scores ranging from approximately 70% to over 90%, depending on task type and level. Qualitative analysis revealed emotional resonance, increased sensitivity to tone, and a perception of authenticity in listening to literary narration. The findings suggest that literary recordings cultivate affective presence, interpretive depth, and intercultural awareness, demonstrating that emotion and cognition can coexist fruitfully within a pedagogy centred on the human voice.

Keywords: Audiobooks; Online English Learning; Literature and Culture; Reflective Learning; Students' Engagement

Resumo

O artigo apresenta um estudo empírico sobre a utilização de audiolivros no ensino *online* de língua inglesa em contexto de ensino superior a distância. Inserido numa reflexão mais ampla sobre a aprendizagem mediada por textos sonoros, o estudo procura compreender de que modo a voz literária pode motivar o empenho, ampliar a interpretação e humanizar a experiência assíncrona. Considera-se também como a voz humana — singular, autêntica e afetivamente densa — pode renovar a empatia e a atenção em espaços digitais cada vez mais mediados pela tecnologia e pela inteligência artificial. Com base em dados recolhidos junto de 528 estudantes adultos de três níveis consecutivos de Inglês, o estudo conjuga registos de participação, indicadores de desempenho e reflexões qualitativas anonimizadas. As sequências de aprendizagem, inspiradas em contos de O. Henry, Chimamanda Adichie e Salman Rushdie, articularam atividades multimodais de escuta, pronúncia e escrita reflexiva. Em todos os níveis, a participação manteve-se estável (71–72%) e as médias de desempenho situaram-se entre aproximadamente 70 e 90%, dependendo do tipo de atividade e nível de domínio da Língua Inglesa. A análise qualitativa revelou ressonância emocional, maior sensibilidade ao tom e

¹ Universidade Aberta, Portugal, Susana.Oliveira@uab.pt, <https://orcid.org/0000-0002-0253-4204>

percepção de autenticidade na escuta da narração literária. Os resultados sugerem que a escuta de registos literários sonoros promove presença afetiva, profundidade interpretativa e consciência intercultural, demonstrando que emoção e cognição podem coexistir de forma produtiva numa pedagogia centrada na voz humana.

Palavras-chave: Audiolivros; Ensino *online* de Inglês; Literatura e Cultura; Aprendizagem Reflexiva; Empenho

1. Introduction

The *Textos Sonoros* project was designed to explore how sound-based texts, particularly podcasts and audiobooks, can enhance language learning in distance education contexts. Implemented by the EI@N research group of the Laboratory of Distance Education and eLearning (LE@D) at Universidade Aberta (UAb), the project analyses the advantages and limitations of these resources, the ways they are used in practice, and the pedagogical principles that should guide the design of sound-based learning sequences.

The English strand of the research group² focuses on enhancing linguistic competence through literary and culturally resonant multimodal experiences centred on three short stories from canonical and contemporary collections: O. Henry's "The Gift of the Magi" (in *The Four Million*, 1906), Chimamanda Ngozi Adichie's "The Thing Around Your Neck" (in *The Thing Around Your Neck*, 2009), and Salman Rushdie's "Good Advice is Rarer than Rubies" (in *East, West*, 1994). Together, these narratives were selected for their thematic significance, contemporary relevance, and stylistic diversity, exposing learners to contrasting voices, registers, and cultural perspectives that reflect the linguistic richness of English and its multiple manifestations within the framework of *World Englishes*.

Within this multilingual and intercultural scope, the English strand also aligns with UAb's pedagogical innovation in distance education. Over the last three decades, distance learning has evolved from analogue correspondence and broadcast media to fully digital ecosystems that sustain a growing, multilayered academic community. Within this trajectory, UAb has developed a mature model of asynchronous, learner-centred education and, as a member of the OpenEU alliance, participates in a network committed to inclusive and sustainable digital transformation across the European Higher Education Area. This institutional framework strengthens UAb's role as a laboratory for innovative pedagogies in e-learning.

It is equally important to recognise the transnational nature of UAb's student body. A significant proportion of learners belongs to the Portuguese-speaking diaspora distributed across multiple time zones. This geographic dispersion does not merely create logistical demands, as it must also be considered when shaping pedagogical choices, assessment practices, and the affective consistency of learning. In this regard,

² The author gratefully acknowledges the contribution of Dr Diana Marques to the development of the *Textos Sonoros* project, particularly in its initial conceptual and pedagogical stages.

UAb is concerned not only with linguistic development but with the formation of interpretive communities capable of listening, interpreting and communicating across distances: social, cultural, and temporal.

In effect, teaching and learning English in asynchronous, fully online environments requires more than linguistic input: it demands strategies that sustain engagement, promote reflection, and connect students emotionally to content. As Byram (2021) argues, linguistic competence and intercultural competence are interdependent dimensions of language education and, as he observes, an emphasis should be put on the “method” used to provide “learners with the means to analyse and thereby understand and relate to whatever social world their interlocutors inhabit” (p. 26).

It is our understanding that the methodological use of audiobooks aligns with this perspective by creating opportunities for learners to engage with authentic voices and narratives that embody diverse cultural contexts.

Within the broader scope of the *Textos Sonoros* project, the English strand study contributes to three of its central objectives: to systematise pedagogical activities that explore the potential of sound-based texts in language learning; to exemplify this exploration through the design of multimodal learning sequences integrating audiobooks and H5P interactive tools in English as a foreign language online teaching; and to test the impact of these sound-based pedagogical paths on selected indicators of learning engagement and performance.

While research has examined digital mediation, multimodality, and affective dimensions of language learning, comparatively less attention has been paid to the pedagogical potential of audiobooks as sustained literary experiences in asynchronous online higher education. In particular, the specific contribution of narrated literary voice to learner engagement, task-based performance, and interpretive awareness in adult distance education remains underexplored.

This study seeks to address this gap by examining how audiobook-mediated sequences, implemented across three CEFR-aligned levels (B1.1 to B2), influence observable participation patterns, measurable learning outcomes, and qualitative dimensions of interpretive development within online English language courses.

The contribution of this article is therefore twofold. Theoretically, it advances a multimodal and humanistic perspective on sound-based literary mediation in online language education. Empirically, it analyses a large dataset drawn from adult distance learners to examine how engagement, performance, and interpretive awareness manifest in audiobook-integrated pedagogical designs.

It is hypothesised that sustained exposure to narrated literary voice in structured asynchronous sequences may enhance participation, support listening-based performance outcomes, and foster deeper interpretive sensitivity to tone, perspective, and intercultural nuance.

Taking these ideas as its point of departure, the present study focuses on three research questions (RQ):

1. How does the integration of audiobooks influence measurable student engagement in asynchronous online English language courses?
2. What relationship can be observed between audiobook-mediated activities and task-based performance outcomes across CEFR levels?
3. In what ways does exposure to narrated literary voice contribute to students' interpretive awareness, particularly in relation to tone, perspective, and intercultural sensitivity?

The following sections outline the conceptual framework underlying these research questions and detail the methodological procedures adopted to operationalise and analyse the study's core constructs.

2. Theoretical Context

At the heart of the English strand *Textos Sonoros* project lies a commitment to enhancing second language acquisition through engagement with Literature and Culture. Byram's (2021) model of intercultural communicative competence provides the theoretical foundation for integrating Literature – or rather “the power of literature” (p. 76), as the author remarks – in language education, as it develops learners' knowledge (*savoirs*), interpretive skills, and attitudes of openness toward cultural difference. In this framework, literary texts are not merely linguistic artefacts but vehicles of cultural identity and critical reflection, offering authentic exposure to vocabulary, syntax, and style.

In support of this argument, Rosenblatt's (1978) transactional theory of reading underlines how meaning emerges through the interaction between the text and the reader — or *listener*. Literature, thus, becomes a space of encounter with the *Other*, where language learning and self-reflection intertwine. Given the global use of English across diverse linguistic and cultural contexts, English Language Teaching (ELT) must therefore cultivate not only communicative competence but also intercultural awareness and sensitivity, valuing cultural diversity as well as human dignity and rights, as Barrett and Borghetti (2025) emphasise in their expansive definition of intercultural competence:

the ability to mobilise and apply psychological resources dynamically in order to respond appropriately, effectively, and respectfully to the demands, challenges, and opportunities that arise within intercultural situations (p. 104).

In this light, including perspectives from multiple world *Englishes* not only reflects the plural nature of the language but also encourages critical engagement with diversity and intercultural awareness. These pedagogical benefits extend to multimodal learning environments, where, as Serafini (2014) points out, combinations of sound, image, and

text invite learners to construct layered meanings, while audiobooks — or “talking literature”, following Rubery’s (2016) observation — democratise access to culture.

Research on audiobooks in English language learning has repeatedly emphasised these resources’ motivational and linguistic benefits (Ayunda, 2013; Padberg-Schmitt, 2020; Lang, M. & Guinovart, X.G., 2021), although empirical evidence in higher education and online settings remains limited. Studies such as Berto and Greggio (2021), Sekścińska and Olszańska (2018) and Saleh et al. (2023) confirm that listening to “talking books” or narrated fiction supports vocabulary development, pronunciation, and affective engagement. Nonetheless, few studies examine these effects in asynchronous or distance-learning environments.

Building upon this growing body of research, studies in distance education (Moore, 1993; Garrison & Anderson, 2003; Dewaele & MacIntyre, 2014; Hilliard et al., 2019; Afonso, Morgado, Carvalho, & Spilker, 2025) have shown that affective, cognitive, and social presence underpin sustained engagement in asynchronous settings. Consequently, the integration of the human voice — at times even the author’s own voice in literary narration — can humanise online interaction and reinforce empathy and attentional focus, as recent studies on emotional design and AI-mediated learning environments have demonstrated (Bozkurt, 2023; Ilieva et al., 2024).

Auditory literary mediation differs from silent reading in that it foregrounds prosody, rhythm, and vocal interpretation as meaning-making resources. The narrator’s voice may function as a scaffold for comprehension, modelling intonation patterns, emotional cues, and pragmatic nuance that are not always fully accessible through text alone. In asynchronous environments, this auditory presence may also contribute to perceived humanisation and relational proximity, potentially influencing motivation and sustained attention.

This intersection between multimodal learning and humanistic education also resonates with Nussbaum’s (2010) and Hakemulder’s (2000) arguments that Literature fosters empathy and ethical imagination. In this sense, the project extends the linguistic, humanistic and intercultural framework and principles to digital pedagogy: encounters with the literary voice become opportunities for language learning and for transformative learning (Mezirow, 1997), supporting both cognitive and emotional growth.

2.1. Conceptual Framework and Research Questions

Research on digital language learning has increasingly emphasised multimodality, affective engagement, and the humanisation of online environments. Studies on online presence (Garrison, 2003), multimodal literacy (Kress, 2010), and affective mediation in language learning (Dörnyei & Ushioda, 2011) suggest that learning in digital contexts benefits from sensory richness and meaningful narrative input. However, while listening comprehension and audio input have long been examined in second language acquisition, the specific pedagogical role of audiobooks as sustained literary experiences in asynchronous higher education remains comparatively underexplored.

Existing research has highlighted the benefits of auditory processing for prosodic awareness, pronunciation modelling, and listening fluency. Yet, fewer studies have examined how narrated literary voice may influence broader dimensions of engagement, task-based performance, and interpretive development in adult distance education. In particular, the potential of audiobooks to foster interpretive sensitivity to tone, narrative perspective, and intercultural nuance has received limited empirical attention.

Within this framework, the present study conceptualises audiobooks not merely as listening materials but as multimodal literary mediators capable of shaping three interrelated domains:

1. Observable engagement in structured online activities;
2. Task-based performance outcomes in listening and analytical writing;
3. The development of interpretive awareness through exposure to narrated voice.

These domains give rise directly to the three research questions guiding this study. RQ1 examines measurable patterns of engagement within audiobook-integrated sequences. RQ2 investigates the relationship between audiobook-mediated tasks and performance outcomes across CEFR levels. RQ3 explores the qualitative dimensions of interpretive awareness emerging from learners' written and reflective responses.

By positioning audiobooks at the intersection of multimodality, affective engagement, and intercultural literary mediation, this study seeks to contribute both conceptual clarity and empirical evidence to an underrepresented area of applied linguistics and online higher education research.

3. Methodology

3.1. Research Design

This inquiry adopts an exploratory empirical design (Bell, Bryman, & Harley, 2022; Creswell & Creswell, 2018), combining descriptive quantitative indicators with qualitative textual analysis. The research aims to understand how sound-based literary materials foster engagement, performance, and interpretive awareness in online English courses. Quantitative data were used to describe participation patterns, while qualitative data provided insight into interpretive and reflective learning.

3.2. Participants

The study was conducted during the 2024/25 academic year, involving UAb's degree learners enrolled in three consecutive English courses.

At the beginning of each course, students were presented with an online survey containing an ethical statement related to the *Textos Sonoros* project. The statement informed participants about the research's objectives, the voluntary nature of participation, and their right to withdraw consent at any time. It also clarified that all data

were collected exclusively for research purposes and processed in compliance with the General Data Protection Regulation (GDPR) and UAb's ethical standards for pedagogical research.

A total of 528 adult learners participated across three consecutive English courses: English I (B1.1), English II (B1.2), and English III (B2). All were enrolled in fully asynchronous environments following the institution's virtual pedagogical model.

3.3. Pedagogical Sequences – Design and Implementation

The implementation of the project followed a pedagogical progression across the three levels. Each learning path was built upon the integration of literary audiobooks, multimodal resources, and reflective writing tasks within the Moodle-based virtual classroom environment. The aim was to explore how the auditory dimension of literature could enhance engagement, comprehension, and interpretive awareness, while fostering oral competence and intercultural reflection in an asynchronous context.

Sequence 1 – “The Gift of the Magi” (English I – Level B1.1)

The first learning path sought to familiarise lower-intermediate learners with literary listening and multimodal comprehension. Centred on O. Henry's “The Gift of the Magi” (1905), it employed the illustrated edition by P. J. Lynch (Candlewick Press, 2008) and a narration by Betsy Bush (LibriVox). The audiobook and written text, based on the Project Gutenberg edition, were embedded in an H5P interactive book hosted within the virtual classroom. The environment was designed as a dedicated learning thematic block, combining audio, text, imagery, and short reflective cues to promote immersion and affective commitment.

Moreover, the virtual classroom privileged a calm visual palette and a measured rhythm of interaction — brief audio segments alternating with reading and reflection activities — so that learners could regulate attention and return to prominent passages at their own pace. Accessibility features, such as adjustable playback speed and captions for selected excerpts, supported inclusion and diverse needs while reinforcing students' sense of control, in accordance with recent research that highlights how inclusion and accessibility in digital higher education require the integration of affective and cognitive dimensions to ensure sustainable participation (Afonso, Morgado, Carvalho, and Spilker, 2025).

It is worth mentioning that the aesthetic and sensory coherence was not meant as ornamental, since it is our understanding that such design considerations and decisions help shape the way learners listened and responded, underlining the interdisciplinary practice and discursive pluralism related to literary meaning.

Thus, integrating image and sound invited a slower, more attentive stance to explore and cultivate thoughtful interpretation and to bring the emotional quality of the text into focus, as the module encouraged participants to explore the emotional undertones of the story

while consolidating vocabulary and sentence structure. Example tasks included brief reflections on tone and meaning, which were followed by discussion forum activities.

Sequence 2 – “The Thing Around Your Neck” (English II – Level B1.2)

The second didactic strand invited learners toward critical listening and intercultural reflection. Chimamanda Ngozi Adichie’s short story “The Thing Around Your Neck” (2009) served as both literary and auditory material, complemented by her TED Talk *The Danger of a Single Story* (2009). Students were asked to analyse tonal shifts, pauses, and emphases, exploring how Adichie’s vocal delivery shaped meaning and emotion.

Within the virtual classroom, a digital mini-dossier structured the learning path into listening comprehension, reflective writing, and thematic analysis. The *Critical Listening* section invited students to write short reflections (40-60 words) on the emotional nuances of the narrator’s voice, while the *Critical Thinking Writing* section encouraged concise essays (100-120 words) addressing narrative technique, identity, and symbolism.

Sequence 3 – “Good Advice is Rarer than Rubies” (English III – Level B2)

The third learning path, implemented in English III (B2), focused on phonetic awareness and accent variation, using the short story as a lexical and thematic context for developing more nuanced pronunciation skills. To this end, three H5P activities — on accent recognition, minimal-pair discrimination, and pronunciation accuracy — guided participants through increasingly complex listening tasks that linked phonological analysis to a meaningful literary framework.

Pedagogically, the module consolidated advanced receptive and productive phonological skills, improving students’ discrimination of segmental features, rhythm, and stress while fostering confidence in oral articulation. Discussions on accent diversity, framed within the world *Englishes* perspective, encouraged learners to value intelligibility and authenticity over imitation of a single variety. Similarly, students also reflected on how their linguistic and cultural backgrounds interacted with the prosodic features of the materials, making the tasks both corrective and intercultural.

3.4. Data Sources

The investigation draws on three primary sources of data collected across the three English courses:

1. Institutional LMS analytics and assessment records, which include completion tracking data for H5P audiobook activities, forum participation logs, and aggregated assessment scores (Critical Listening and writing tasks).
2. Qualitative data were derived from open-ended written students’ reflective responses in mini-dossiers, written analytical submissions, and forum contributions linked to the audiobook sequences.

3. Selected perception items included in course questionnaires captured learners' views on linguistic accessibility, perceived support from narration, and the relevance of sound-based tasks. These items formed part of the standard course questionnaires and were structured as 5-point Likert-scale statements (ranging from strongly disagree to strongly agree), and response distributions were analysed at cohort level to contextualise reported engagement patterns.

All data were anonymised and analysed at cohort level.

Table 1
Example of Aggregated Participation Data

Course	No. of enrolled learners	Number of Participants	Completion rate (%)	Mean scores (%)
English I (B1.1)	208	150	72	91
English II (B1.2)	173	123	71	70.5 / 63.25
English III (B2)	147	106	72	79

Source: Universidade Aberta's institutional analytics reports (LMS), 2024/25.

The ensuing table illustrates the adopted framework, enabling instructors to systematically collect qualitative evidence and connect it to the project's research questions.

Table 2
Qualitative Evidence Framework for Instructor Data Collection (English II – B1.2)

Mini-Dossier section/Question focus	Type of evidence to collect (qualitative dimension)	Connection to Research Questions (RQ)	Illustrative examples of what to record
Part 2 Critical Listening Writing	Learner reflections on tone, pronunciation, rhythm, pace, intonation, and emotional effects of voice	RQ1: Engagement (attention to authentic voice and listening motivation) RQ3: Affective/interpretive perception	Comments describing how voice conveys irony, distance, surprise, sadness, or empathy (e.g., "The narrator sounds ironic when describing her family's expectations")
Audio Excerpt 1: Family's expectations	Awareness of irony, detachment, or affection in the narrator's tone	RQ3	Notes on learners perceiving sarcasm or warmth; recognition of emotional stance toward cultural stereotypes
Audio Excerpt 2: Surprise about the university system	Recognition of tone shifts, hesitation, pace changes expressing emotion	RQ1 & RQ3	Observations linking tone or rhythm to frustration, curiosity, or confusion
Audio Excerpt 3: Blank fortune cookie	Interpretation of intonation/emphasis and emotional impact (disappointment, resignation)	RQ3	Learner references to sadness or symbolic silence; empathy with narrator

Source: English strand research-team framework for instructor qualitative data collection (*Textos Sonoros* project, Universidade Aberta, 2024/25).

Underpinning these findings is the assumption that listening is not a passive act but a cognitively and emotionally layered process.

3.5. Key Constructs, Measures, and Analytical Procedures

The study examines three central constructs: engagement, performance, and interpretive awareness. Engagement was operationalised as behavioural engagement (completion and participation data) and self-reported affective engagement (perception items).

For example, in English II, 123 out of 173 enrolled students completed the audiobook-based sequence (71%). Similar completion patterns were observed across the other courses (see Table 1). Perception items were analysed descriptively to contextualise learners' reported engagement with narrated literary materials.

Performance refers to task-based learning outcomes associated with audiobook-mediated activities. Critical Listening scores were recorded on a 0–1 scale and reported as percentage equivalents (e.g., $0.705/1.0 = 70.5\%$). Writing performance was evaluated using a structured rubric distinguishing four performance bands (Excellent, Good, Satisfactory, Limited), based on interpretive precision, thematic development, structural coherence, and linguistic control.

In English II, among the 123 students who completed the activity, the mean listening score was $0.705/1.0$ (70.5%), and the average writing performance reached 63.25% according to the established rubric criteria.

No inferential statistical modelling was conducted; descriptive statistics were used to identify stable patterns across cohorts.

Interpretive awareness was conceptualised as students' demonstrated capacity to recognise tone, narrative voice, perspective, and intercultural nuance in literary texts presented in audio form.

Qualitative data from mini-dossiers, analytical responses, and forum posts were analysed using a construct-informed thematic approach. Initial coding categories were derived deductively from the study's core constructs, while allowing inductively emerging subthemes to surface through successive rounds of close reading, comparison across responses, and collaborative refinement of categories.

Coding was conducted collaboratively by the two principal lecturers responsible for course design and assessment, with subsequent review and discussion involving course tutors. All reflective submissions associated with the audiobook sequences were included in the analysis at cohort level, ensuring comprehensive coverage of the qualitative dataset rather than selective sampling. This iterative analytical process supported category refinement and interpretive coherence across data sources. Discrepancies in coding were resolved through discussion and consensus among the research team.

As an exploratory study situated within pedagogical practice, the analysis focuses on identifying consistent participation and performance patterns across cohorts. Broadly speaking, the responses indicate sustained participation patterns alongside qualitatively inferred affective engagement, as reflected in learners' thematic references to tone, emotion, and narrative presence.

3.6. Data Analysis

Following the implementation of the pedagogical sequences described above, data were analysed using a mixed-methods approach, combining quantitative and qualitative procedures and addressing the three research questions guiding the present *Textos Sonoros* project study.

Consistent with its exploratory design, the analysis maps participation tendencies and performance patterns without advancing claims of causality.

Quantitative data from the LMS were analysed descriptively to identify trends in participation, completion, and performance. Mean scores, number of attempts, and participation rates were compared across the three English courses to reveal overall engagement and achievement patterns.

These indicators also provided the basis for addressing RQ2, finding evidence to support measurable learning outcomes associated with audiobook-based activities.

Overall, the consistency observed in participation rates (71–72%), together with mean performance scores ranging from 70% to 91%, suggests a steady level of engagement and measurable progress across courses and activity types.

In the first two didactic strands, which were complemented by discussion forums, learners' posts further illustrated behavioural engagement and peer interaction, complementing the evidence drawn from activity completion and performance data.

Concerning qualitative data from the reflective tasks, these were analysed thematically. Following Braun and Clarke (2006, 2014) and Nowell et al. (2017), instructor summaries were read iteratively to identify recurring ideas, which were then grouped into broad categories capturing affective and interpretive dimensions of learning identified in students' reflections.

Table 3
Illustrative Categories from Thematic Analysis (English II)

Category	Description	Representative Learner Observation
Awareness of tone	Recognition of how intonation and voice convey emotion and meaning	"Listening to the narrator helped me realise how irony changes the message."
Emotional resonance	Personal connection to character emotions or situations	"I felt moved by the narrator's calm tone; it made the story more powerful."
Intercultural reflection	Awareness of cultural perspectives and empathy for other contexts	"The story made me compare my own background with the narrator's."

Source: English II (B1.2) instructor-compiled reflections (Universidade Aberta, 2024/25).

The integration of descriptive statistics and thematic analysis enabled a balanced understanding of engagement, measurable learning outcomes, and interpretive development.

4. Results and Discussion

The data gathered across the three English courses confirm that sound-based literary activities can sustain both engagement and learning in asynchronous higher education. Participation remained remarkably consistent — 72% in English I, 71% in English II, and 72% in English III — providing evidence of sustained behavioural engagement among adult learners scattered across different geographies and life circumstances. Within the particular characteristics of the UAb's student body and pedagogical model, these results should not be read merely as compliance but as evidence of sustained behavioural engagement.

4.1. Engagement and Affective Presence

At the introductory level, most participants described the sound-based text “The Gift of the Magi” as linguistically accessible (82%) and the narrator's tone as supportive of comprehension (68%).

Forum participation, on the other hand, was modest, with only 18% of the cohort sharing reflections with peers and instructors. Although such a result may align with the ethos of asynchronous learning in general, the pattern observed here is lower than what is usually found in UAb's foreign-language courses. The fact that this activity coincided with the week preceding Christmas may partly explain the outcome, as this period is typically constrained by adult learners' professional and family commitments. Nevertheless, the short-story activities were deliberately timed to resonate with the Christmas season and to promote reflection on family, generosity, and love, values that acquire renewed significance at this time of year.

In this context, it is worth recalling that affective and cognitive presence, as Moore (1993) and Garrison and Anderson (2003) have long suggested, are more meaningful indicators of engagement than numerical participation alone.

Be that as it may, listening to O. Henry's narrative during the week preceding Christmas invited emotional proximity and moral reflection. Many participants wrote about generosity and sacrifice in relation to contemporary forms of labour and affection, transforming an act of listening into an act of recognition. The optional H5P pronunciation task — completed by 42% of participants — further revealed autonomy and affective investment, as students sought to hear and rehear their own voices within the shared literary space.

4.2. Learning and Performance

In “The Thing Around Your Neck” (English II), the sound-based sequence featured the author’s own voice — Adichie reading her short story — and the pedagogical design advanced toward critical listening activities.

Out of 173 enrolled students, 123 completed the audiobook-based sequence (71%). Among those who submitted the tasks, the mean Critical Listening score was 0.705/1.0 (70.5%), and the average writing performance reached 63.25% according to the established rubric-based assessment criteria.

Learners who completed the audiobook-based tasks and demonstrated sustained listening engagement frequently produced richer analytical writing, suggesting that sustained listening sharpened textual awareness and interpretive precision. As one participant noted, “Her tone sounds ironic and slightly amused, suggesting distance from stereotypes.” Another observed that Adichie’s “lower pitch conveys disappointment and quiet acceptance.” Together, these comments illustrate the emergence of what might be called *interpretive listening*, a concept aligned with the development of interpretive awareness as defined in Section 3.5: the ability to connect tone, rhythm, and irony with emotion, identity, and worldview.

If this interpretation holds, the findings resonate with recent research showing that multimodal learning fosters deeper cognitive processing through the integration of auditory and verbal input (Li, 2025; Berto & Greggio, 2021). In this sense, listening ceases to be a passive skill and becomes a cognitive act of reading through sound.

4.3. Phonetic Awareness and Cultural Listening

By the advanced level (“Good Advice is Rarer than Rubies”), learners had moved from interpretive to productive engagement. The three H5P activities combined accent recognition, minimal-pair discrimination, and pronunciation analysis, integrating literary content with phonetic precision. Average scores ranged between 68.6% and 86.4%, with a global mean of approximately 79%, confirming the potential of literary materials to consolidate oral accuracy in distance learning contexts. These percentages correspond to the three distinct H5P phonological activities described in Section 3.3.

Beyond technical ability, however, students reflected on accent diversity and intelligibility, recognising English as a plural and global phenomenon. Discussions often centred on the idea that intelligibility — rather than imitation of a single prestige variety — defines communicative competence. This perspective is consistent with the world *Englishes* framework and particularly relevant in multilingual academic environments and multicultural societies.

In short, the progression across levels delineated a coherent pedagogical arc: from affective resonance to critical reflection, and from comprehension to oral self-expression. Listening thus evolves into reading, reflection, and ultimately into voice, both literal and metaphorical.

Table 4

Comparative Overview of Participation, Performance, and Affective Indicators across the Three Audiobook-Based Sequences (English I–III)

Sequence	Literary work	Participation	Mean score (listening/writing)	Affective indicators	Key themes
I	“The Gift of the Magi” (O. Henry)	72% (150/208)	91% (formative H5P average)	82% language accessibility; 68% narrator supports comprehension	Generosity; resilience
II	“The Thing Around Your Neck” (C. Adichie)	71% (123/173)	70.5% / 63.25%	74% tone awareness; 69% writing clarity	Identity; belonging
III	Good Advice is Rarer than Rubies (S. Rushdie)	72% (106/147)	68.6% / 86.4% / 83.1% (mean ≈ 79%)	81% perceived relevance; 66% confidence in pronunciation	Accent diversity; intelligibility

Source: Aggregated LMS analytics and course-embedded perception items (2024/25).

4.4. Discussion

The integrated results suggest that emotion, reflection, and autonomy are not peripheral to online asynchronous learning but should be regarded as some of its sustaining forces. Emotional connection operates through cognitive and interpretive processes, enabling learners to translate empathy into awareness and awareness into agency. As Saleh et al. (2023) argue, affective engagement promotes achievement through self-regulation and organisation of knowledge, an observation corroborated by these findings.

Equally significant is the cultural dimension of listening. The audiobooks introduced distinct worldviews — O. Henry’s moral realism, Adichie’s postcolonial irony, and Rushdie’s hybrid cosmopolitanism — inviting participants to inhabit other voices and to recognise their own within them. In a context increasingly mediated by artificial intelligence and standardised communication, whether written or spoken, such encounters reaffirm the pedagogical value of the authentic human voice as a bridge between emotion and cognition (Bozkurt, 2023).

These findings may also be read in light of Rosenblatt’s (1978) transactional theory, according to which meaning emerges through the dynamic interaction between reader and text. In the present study, narrated literary voice appears to mediate this transaction by foregrounding tone, rhythm, and affective cues that shape interpretation. Similarly, Byram’s (2021) model of intercultural communicative competence helps frame the interpretive movement observed in learners’ responses: listening becomes not only comprehension but relational engagement with cultural perspectives embodied in voice.

Taken together, Sections 4.1–4.3 address the three research questions by showing stable participation patterns (RQ1), consistent task-based performance outcomes across levels (RQ2), and qualitative evidence of interpretive awareness linked to tone, perspective, and intercultural sensitivity (RQ3).

In sum, the present study demonstrates that affective engagement, interpretive reflection, and cognitive precision can coexist productively in online asynchronous foreign-language learning. These findings reinforce the theoretical alignment between multimodal mediation, intercultural communicative competence, and affective presence outlined in Section 2. Emotion here is not a decorative element but a pathway, one that leads from listening to understanding, from comprehension to interpretation, and finally from reflection to agency.

5. Limitations and Future Directions

The investigation draws on institutional analytics and instructor-compiled qualitative evidence across three consecutive English courses within one institution. While the sample is large for an exploratory design, participation was voluntary and self-selected, and no control group was used. Sociodemographic variables were intentionally not collected to align the scope with pedagogical aims. Although collaborative coding enhanced analytical robustness, future research could incorporate formal inter-coder reliability procedures and triangulated interview data to further strengthen validity. Future work should also aim at triangulating instructor records with learner interviews, including additional cohorts and institutions to compare spoken literature-based tasks with functionally equivalent non-audio conditions in order to isolate modality effects.

6. Conclusion and Implications

The present empirical inquiry demonstrates that affective engagement, interpretive reflection, and digital literacy can coexist harmoniously in online asynchronous foreign-language education, showing that integrating audiobooks within virtual classrooms can promote sustained attention and self-regulated learning. By merging human voice, literary text, and interactive design, the project reintroduces the Humanities as a living component of e-learning.

Beyond measurable outcomes, the study underscores the role of emotional well-being as a driver of cognition: enjoyment and balance are not merely desirable but essential mechanisms for motivation, persistence, and success in online learning (Runa, Oliveira & Cardoso, 2022; Oliveira & Runa, 2021). When learners listen to a narrator's voice, they engage in a form of dialogue that transcends distance and bridges the cognitive and affective dimensions of understanding.

The proposed model is transferable across languages and disciplines, provided that audio-mediated materials are carefully selected and embedded in reflective pathways that move from listening to writing and, where feasible, to speaking. Adaptations should attend explicitly to accent diversity and to the ethics of representation in transnational cohorts. In this sense, *Textos Sonoros* complements inclusive design frameworks by

foregrounding human voice and interpretive agency as levers for engagement and well-being.

In future work, the project aims to explore the integration of sound and Literature in multilingual contexts and to develop comparative models across linguistic and cultural systems. Ultimately, learning through sound in distance education invites a broader reflection on how the Humanities can foster empathy, creativity, and ethical awareness in the digital age.

References

- Adichie, C. N. (2009). *The thing around your neck*. Alfred A. Knopf.
- Afonso, A., Morgado, L., Carvalho, I. C., & Spilker, M. J. (2025). Facing Challenges in Higher Education: Enhancing Accessibility and Inclusion Through Flexible Learning Design. *Education Sciences*, 15(8), 1013. <https://doi.org/10.3390/educsci15081013>
- Ayunda, A. (2013). The effect of audiobook use on EFL students' fluency development. *Journal on English as a Foreign Language*, 3(2), 85–92. <https://doi.org/10.23971/jefl.v3i2.67>
- Barrett, M. & Borghetti, C. (2025). What Is Intercultural Competence?. In P. Genkova, M. Flynn, M. Morley, & M. Rašticová (Eds.), *Handbook of Diversity Competence* (pp.93–108). Springer. https://doi.org/10.1007/978-3-031-69308-3_7
- Bell, E., Bryman, A., & Harley, B. (2022). *Business Research Methods* (6th ed.). Oxford University Press.
- Berto, E.F. & Greggio, S. (2021). As Potencialidades do Gênero Podcast no Desenvolvimento e Aprimoramento da Habilidade de Compreensão Oral na Aprendizagem de Língua Inglesa. *Ilha do Desterro: Revista de Língua Inglesa, Literaturas em Inglês e Estudos Culturais*, 74(3). <https://doi.org/10.5007/2175-8026.2021.e79454>
- Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2014). What Can "Thematic Analysis" Offer Health and Wellbeing Researchers? *International journal of qualitative studies on health and well-being*, 9, 26152. <https://doi.org/10.3402/qhw.v9.26152>
- Byram, M. (2021). *Teaching and assessing intercultural communicative competence: Revisited* (2nd ed.). Multilingual Matters. <https://doi.org/10.21832/9781800410251>
- Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education*, 18(1), 198–204. <https://doi.org/10.5281/zenodo.7716416>

- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage.
- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237–274. <https://doi.org/10.14746/ssl.2014.4.2.5>
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Pearson.
- Garrison, D. R., & Anderson, T. (2003). *E-Learning in the 21st Century: A Framework for Research and Practice*. Routledge.
- Hakemulder, J. (2000). *The Moral Laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept*. John Benjamins.
- Henry, O. (1906). *The four million*. Doubleday, Page & Company.
- Hilliard, J., Donelan, H. & Heaney, C. (2019). Exploring the Emotions of Distance Learning Students in an Assessed, online, Collaborative Project. In *EDEN Conference Proceedings* (pp. 260–268). EDEN. <https://doi.org/10.38069/edenconf-2019-ac-0028>
- Ilieva, G., Yankova, T., Ruseva, M., & Kabaivanov, S. (2025). A Framework for Generative AI-Driven Assessment in Higher Education. *Information*, 16(6), 472. <https://doi.org/10.3390/info16060472>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Lang, M. & Guinovart, X.G. (2021). Developing and implementing an English-Spanish literary parallel audio-textual corpus for data-driven ESL learning. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 37(1), 2021370106. <https://doi.org/10.1590/1678-460X2021370106>
- Li, M. (2025). Impact of multimodal learning environments on cognitive and emotional development in students. *Journal of Computational Methods in Sciences and Engineering*. <https://doi.org/10.1177/14727978251363052>
- Mezirow, J. (1997). Transformative learning: Theory to practice. *New Directions for Adult and Continuing Education*, 74, 5–12. <https://doi.org/10.1002/ace.7401>
- Moore, M. (1993). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical Principles of Distance Education* (pp. 22–38). Routledge.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, 16(1). <https://doi.org/10.1177/1609406917733847>
- Nussbaum, M. C. (2010). *Not for Profit: Why Democracy Needs the Humanities*. Princeton University Press.

- Padberg-Schmitt, B. (2020). Increasing Reading Fluency in Young Adult Readers Using Audiobooks. *CLELE Journal*, 8(1), 31–51.
- Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Southern Illinois University Press.
- Oliveira, S., & Runa, A. (2021). To teach and delight: Emotions and second language acquisition onsite and online practices. In *Proceedings of the iTEL 2021 – International Conference on Innovation and Technology in Language Education* (pp. 45–58). Universidade Aberta.
- Rubery, M. (2016). *The untold story of the talking book*. Harvard University Press.
- Runa, A., Oliveira, S., & Cardoso, C. (2022). *Emotions, Online Teaching and Online Evaluation in Higher Education during a Pandemic*. *RE@D – Revista de Educação a Distância e eLearning*, 5(2), 111–131.
<https://doi.org/10.34627/redvol5iss2e202209>
- Rushdie, S. (1994). *East, West*. Jonathan Cape.
- Saleh, S., AlAli, R., Wardat, Y., Al-Qahtani, M., Soliman, Y., & Helali, M. (2023). Structural Relationships between Learning Emotion and Knowledge Organization and Management Processes in Distance Learning Environments: “An Applied Study”. *European Journal of Investigation in Health, Psychology and Education*, 13(9), 1569-1589. <https://doi.org/10.3390/ejihpe13090114>
- Sekścińska, I., & Olszańska, M. (2018). The Importance of Audiobooks in English Language Teaching. *Polish Journal of Applied Sciences*, 4(4), 143–149.
<https://doi.org/10.34668/PJAS.2018.4.4.04>
- Serafini, F. (2014). *Reading the Visual: An Introduction to Teaching Multimodal Literacy*. Teachers College Press.

Recebido 27/10/2025
Aceite 11/03/2026
Publicado 07/04/2026

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)
