



Enhancing Podcast Accessibility: The Role of Audio Descriptions for Visually Impaired Listeners

Dr. Ibrahim Abdelfattah Almajali

Faculty of art's, Arabic department, King Faisal University, Alhafof, The Eastern Province,
Saudi Arabia.

ialmajali@kfu.edu.sa

Abstract:

As podcast consumption continues to grow globally, ensuring equitable access for all listeners, including those with visual impairments, has become a critical concern. Podcasts predominantly rely on auditory storytelling but often include implicit visual or contextual elements that may be inaccessible to visually impaired audiences. This paper investigates the role of audio descriptions—narrative elements that verbally convey visual information—in enhancing podcast accessibility for visually impaired listeners. While audio description has traditionally been associated with film and television (Gustavsson, 2019; Remael et al., 2021), its application in podcasting remains underexplored. Through a mixed-methods approach combining a systematic literature review, usability testing with visually impaired podcast users, and a technical analysis of audio description integration techniques, this study identifies best practices and challenges in adapting audio descriptions for audio-only media.

The findings reveal that embedding descriptive narration about visual cues, speaker gestures, background sounds, and non-verbal communication improves listener comprehension and engagement without detracting from the storytelling experience (Barton & Coughlan, 2022). Additionally, user feedback highlights the importance of customizable description levels to cater to individual preferences and contexts (Lee et al., 2023). The paper also addresses technological considerations, such as synchronization of descriptions with podcast timelines and the use of metadata standards for accessible audio content (W3C, 2024).

This research contributes to the growing discourse on media accessibility by proposing a framework for implementing audio descriptions in podcasts, advocating for inclusive production guidelines, and encouraging collaboration between content creators, accessibility experts, and platform developers. Ultimately, this work supports the goal of making podcasts a fully inclusive medium, aligning with international accessibility mandates such as the United Nations Convention on the Rights of Persons with Disabilities (UN CRPD, 2006) and advancing digital equity.



Keywords: Podcast Accessibility, Audio Description, Visually Impaired, Listeners Inclusive, Media Accessible, Audio Content.

Introduction

Podcasts have rapidly become one of the most widely consumed forms of media, with global listeners expected to surpass 500 million in 2024 (Statista, 2024). Although podcasting is inherently an auditory medium, many episodes include visual or context-dependent elements such as references to on-screen actions, gestures, facial expressions, or images. For listeners who are blind or visually impaired, this creates a significant barrier to full comprehension and enjoyment (Lopes et al., 2023).

Audio Description (AD), a technique originally developed for film and television, narrates visual details that are essential to understanding the narrative (Remael et al., 2021). AD is increasingly recognized as a vital component of accessible media, endorsed by standards such as the Web Content Accessibility Guidelines (WCAG) and initiatives by the World Wide Web Consortium (W3C, 2023). However, little research has been done on adapting audio descriptions for audio-only formats like podcasts.

Recent studies highlight the potential of adaptive and user-controlled audio descriptions. Lee et al. (2024) developed a context-aware, AI-driven system that tailors descriptions to listener preferences, improving engagement and comprehension. Similarly, Natalie et al. (2024) introduced customizable AD tools that allow users to adjust voice, pacing, and emotional tone to match personal needs. These systems demonstrate that audio descriptions are not only beneficial but can be made flexible and non-intrusive.

In addition to enhancing accessibility for blind and visually impaired listeners, AD benefits a wider audience, including neurodiverse individuals, language learners, and multitaskers who rely on rich auditory cues (Wang & Taele, 2025). Podcasts, therefore, present a compelling opportunity to reimagine AD not as an afterthought, but as a core narrative component.

This study explores the integration of audio descriptions into podcast production, seeking to:

1. Adapt existing AD frameworks for use in audio-only storytelling;
2. Evaluate the effectiveness of descriptions in improving listener comprehension, immersion, and satisfaction;
3. Offer practical recommendations for podcasters and content creators to implement accessible design.



By bridging the gap between accessibility research and audio storytelling, this work contributes to more inclusive media experiences and supports the broader goal of digital equity for people with disabilities.

Literature Review

1. Foundations of Audio Description (AD) in Media

Audio Description (AD) is a core accessibility technique developed to make visual content available to blind and visually impaired audiences by narrating essential visual elements during natural pauses in dialogue. Originally adopted in film and television, AD has evolved from static, objective narration to more creative and immersive forms (Lopez et al., 2022; Remael et al., 2021). Studies have shown that AD improves narrative comprehension, spatial awareness, and enjoyment among blind users (Fryer & Freeman, 2013; Hättich & Schweizer, 2020).

Despite its benefits, traditional AD often struggles to integrate smoothly with creative soundtracks or complex visual content. This has led researchers and developers to explore more dynamic forms of AD that are both informative and engaging (Lopez et al., 2022).

2. Innovations in AD Craftsmanship

Lopez et al. (2022) introduced “enhanced audio description,” incorporating binaural sound, first-person narration, and immersive sound design to blend AD into the narrative more naturally. Tools like *Rescribe*, developed by Pavel et al. (2020), have enabled content creators to automatically identify and fit AD segments into pauses without manual segmentation. These approaches allow AD to support both accessibility and creative storytelling.

3. Automating and Personalizing Audio Description

With advances in artificial intelligence, new systems are now automating AD production. Lee et al. (2024) developed a context-aware, AI-based system (*CA³D*) that detects key moments for description and inserts adaptive narration. Similarly, Gao et al. (2024) explored the use of vision-language models and large language models (LLMs) to automate and personalize audio description. These models allow users to control the tone, speed, and verbosity of AD, enhancing the listening experience and broadening its appeal beyond the visually impaired community.



4. Accessibility in Audio-First Contexts

While AD is well-established in visual media, its application to audio-first formats like podcasts remains minimal. Studies show that even in auditory environments like Clubhouse, visually impaired users face usability challenges due to inaccessible UI components and lack of descriptive cues (Rosenberg et al., 2024). Tools and workflows for podcast creators that support screen readers, keyboard navigation, and AD authoring are still underdeveloped, highlighting a critical area for future research and development.

5. Broader Inclusive and Educational Benefits

The impact of AD extends beyond visually impaired users. It benefits second-language learners, neurodiverse individuals, and those who rely on audio content while multitasking (Wang & Taele, 2025). In educational settings, AD supports comprehension and engagement by providing additional context through narrative, especially in STEM and media literacy courses (Hersh, 2023). Moreover, the DAISY Consortium has shown how structured audio and navigation features can offer a more empowering experience for visually impaired users in both educational and recreational contexts (DAISY, 2023).

Research Method

1. Research Design

This study employs a **mixed-methods approach**, combining qualitative and quantitative techniques to explore how audio description (AD) can be effectively integrated into podcasts for visually impaired listeners. A multi-phase design was adopted:

1. **Systematic Literature Review** – Identifying existing AD and audio accessibility frameworks.
2. **Focus Groups & Interviews** – Eliciting firsthand experiences and preferences from visually impaired podcast users.
3. **Pilot Implementation & Testing** – Producing AD-enhanced podcast segments and evaluating impact on comprehension and engagement.

2. Literature Review

A comprehensive review was conducted following Wolfswinkel et al.'s (2013) six-phase protocol for systematic reviews applied to AD and accessibility studies. Core sources included research on AD in film (Perego, 2018; Fresno, 2022), automated description tools (Pavel et al., 2020), and context-aware systems (Lee et al., 2024).



3. Participants

For user testing, we recruited:

- **12 visually impaired adults** (6 male, 6 female; ages 18–65) via local blindness associations.
- **6 podcast producers** (3 with experience in AD, 3 without) via industry outreach.

4. Data Collection Methods

a. Focus Groups & Interviews

Three focus groups (4 participants each) and six semi-structured interviews (30–45 minutes) gathered participant insights on:

- Challenges with visual references in podcasts.
- Desired AD features (length, tone, timing).
- Integration preferences within episodic storytelling.

Data was audio-recorded, transcribed, and thematically coded following Braun and Clarke (2021).

b. Pilot Podcast Segments

Three 10-minute podcast excerpts were produced:

- **Transcript-only** version.
- **Traditional AD** version.
- **Enhanced AD** version with immersion features (binaural cues and emotional tone), informed by López et al. (2022).

Segments were reviewed and refined using Rescribe (Pavel et al., 2020) to optimize AD placement.

5. Evaluation Procedures

Participants listened to all three segments in randomized order. After each, they completed:

- A **comprehension quiz** (10 multiple-choice items).



- A **10-point Likert questionnaire** assessing narrative presence, enjoyment, and perceived accessibility.

Paired t-tests compared comprehension and satisfaction scores between versions.

6. Data Analysis

- **Qualitative** interviews were coded inductively and analyzed for emerging themes.
- **Quantitative** data were analyzed using SPSS, assessing differences via paired t-tests and effect sizes.
- **Mixed-method integration** was conducted using a convergence design to align qualitative insights with quantitative outcomes.

Results and Discussion

1. Quantitative Results

Twelve visually impaired participants evaluated three podcast versions: (1) Transcript-only, (2) Traditional Audio Description (TAD), and (3) Enhanced Audio Description (EAD). The results showed clear differences in comprehension and user satisfaction:

- **Comprehension Scores (out of 10):**
 - Transcript-only: $M = 6.4$, $SD = 1.3$
 - TAD: $M = 8.1$, $SD = 1.1$
 - EAD: $M = 9.2$, $SD = 0.6$
- **Engagement Ratings (10-point scale):**
 - Transcript-only: $M = 5.8$
 - TAD: $M = 7.6$
 - EAD: $M = 9.1$

Paired-sample t-tests confirmed statistically significant differences between the transcript-only and AD-enhanced formats ($p < .01$). The enhanced AD version significantly outperformed the traditional version in both comprehension and enjoyment, echoing findings by López et al. (2022) that immersive audio features (e.g., first-person perspective, binaural effects) increase listener involvement.



2. Qualitative Themes from Interviews and Focus Groups

a. Desire for Integration, Not Interruption

Participants consistently preferred audio descriptions that blended seamlessly with the narrative rather than those inserted rigidly between dialogue. They described traditional AD as “too clinical” or “mechanical,” while the enhanced AD was described as “natural” and “part of the story.”

“The second one [enhanced AD] felt like it was written with the podcast, not added afterward.”

(Participant 6)

This aligns with enhanced AD research suggesting better narrative cohesion leads to deeper cognitive immersion (Fresno, 2022).

b. Importance of Emotional Tone

Participants valued AD that reflected the tone of the episode. Emotionally aware narration helped them feel more connected to the content and characters.

“When the voice matched the emotion—like fear or sadness—I could follow the story better, not just what was happening, but how it felt.”

(Participant 2)

This supports findings by Gao et al. (2024), who emphasized the value of emotionally adaptive AD in immersive environments.

c. Control and Customization

Several users expressed interest in being able to toggle or customize the level of description, supporting the growing emphasis on user-personalized AD systems (Natalie et al., 2024; Lee et al., 2024). Some preferred minimal description for faster pacing, while others appreciated more detailed descriptions during complex scenes.

3. Challenges and Production Implications

While enhanced AD received higher satisfaction scores, podcast producers in the study noted its higher production complexity. Synchronizing narration, preserving pacing, and mixing immersive elements require advanced audio engineering skills and accessible tools such as



Rescribe (Pavel et al., 2020). Producers without AD experience expressed concern over time constraints and lack of guidance on AD best practices.

This gap highlights the need for **standardized guidelines** for podcast AD, similar to WCAG or the W3C's Accessible Media Guidelines (W3C, 2023), and easy-to-use authoring tools.

4. Synthesis and Implications

This study provides empirical support for integrating audio description into podcast production. Enhanced AD significantly improves comprehension, emotional engagement, and overall satisfaction for blind and visually impaired users. The findings reinforce that AD should not be treated merely as a compliance measure but as an inclusive storytelling tool that benefits a wider range of users, including those with neurodiversity, limited literacy, or auditory processing challenges (Wang & Taele, 2025).

Moreover, the demand for customization echoes broader trends in user-centered accessibility, emphasizing flexibility and personalization as core design principles.

Conclusion

This study demonstrates the significant impact of integrating audio description (AD) into podcast production to improve accessibility for visually impaired listeners. Through both quantitative assessments and qualitative feedback, it was evident that audio descriptions—especially those designed with immersive, emotional, and narrative-aware elements—substantially enhance listener comprehension, engagement, and satisfaction.

Participants responded most positively to enhanced AD formats that felt organically woven into the storytelling, highlighting the importance of creative and context-sensitive implementation. Additionally, the desire for customizable AD features underscores a growing expectation for user-centered accessibility design, not only among blind users but across broader audiences with diverse needs.

While the adoption of AD in podcasts is still in its early stages, this research confirms its feasibility and value. However, it also brings attention to the production challenges faced by podcast creators, including a lack of accessible tools, standardized practices, and training resources. Addressing these barriers will be crucial in fostering inclusive podcast ecosystems.



Ultimately, this study advocates for a shift in how audio accessibility is approached—not as an afterthought or a legal obligation, but as a means of enriching the storytelling experience for all listeners.

Recommendations for Future Study

Building on the findings of this research, several avenues for future investigation are recommended to further enhance podcast accessibility through audio description:

1. **Development of User-Centered Tools:** Future studies should focus on creating and testing accessible, intuitive tools that enable podcast creators—especially those without audio engineering expertise—to produce high-quality audio descriptions easily and efficiently.
2. **Exploration of Personalization Features:** Research is needed to better understand how customizable audio description settings, such as pacing, detail level, and voice tone, affect different user groups, including neurodiverse listeners and second-language users.
3. **Longitudinal Impact Studies:** Long-term studies tracking how repeated exposure to audio-described podcasts influences listening habits, comprehension skills, and overall engagement among visually impaired audiences would provide valuable insights.
4. **Broader Audience Inclusion:** Future work should explore how audio description can benefit other populations, such as individuals with cognitive or learning disabilities, as well as those using podcasts in multitasking scenarios.
5. **Integration with Emerging Technologies:** Investigating the use of artificial intelligence, machine learning, and natural language processing for semi-automated or fully automated audio description in podcasting holds promise for scalable solutions.
6. **Accessibility Standards and Guidelines:** Further research could contribute to establishing industry-wide standards and best practices tailored specifically to audio description in podcasting to guide producers and platform developers.

By addressing these areas, future research can contribute to more inclusive audio media environments that embrace diverse listener needs and preferences.

conflict of interest

The author declare no conflict of interest



Funding : This work was supported by the Deanship of Scientific Research, King Faisal University, Saudi Arabia Grant No.KFU252436

References :

- [1] Barton, L., & Coughlan, T. (2022). Enhancing audio narratives: The role of audio description in storytelling media. *Journal of Accessibility Studies*, 15(3), 145–162.
- [2] Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications.
- [3] DAISY Consortium. (2023). Accessible publishing best practices.
- [4] Fresno, N. (2022). Research in audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of audio description* (pp. 312–327). Routledge.
- [5] Fryer, L., & Freeman, J. (2013). Cinematic language and the description of film: Keeping AD users in the frame. *Perspectives: Studies in Translatology*, 21(3), 412–426.
- [6] Gao, H., Xu, Y., Li, T., & Wang, M. (2024). Adaptive multimodal audio description using large language models. *Journal of Artificial Intelligence for Media Accessibility*, 2(1), 22–35.
- [7] Gustavsson, M. (2019). Audio description in the age of streaming media. *Accessibility and Disability Review*, 8(2), 90–104.
- [8] Hättich, A., & Schweizer, K. (2020). Listening to what cannot be seen: The cognitive processing of audio description. *British Journal of Visual Impairment*, 38(2), 145–156.
- [9] Hersh, M. A. (2023). Inclusive audio-based learning in higher education: Audio description and podcasting. *Journal of Educational Technology & Society*, 26(1), 40–53.
- [10] Lee, S., Kim, H., & Park, J. (2023). User preferences in audio description for accessible podcasts. *International Journal of Human-Computer Interaction*, 39(4), 321–337.
- [11] Lee, S.-H., Wang, J., Fan, D., Zhang, Z., Liu, L., Hao, X., Bhat, V., & Li, X. (2024). NowYouSeeMe: Context-aware automatic audio description. *arXiv*.
- [12] Lopes, R., Marques, C., & Maia, R. (2023). Enhancing podcast accessibility: Designing audio for all. *Journal of Accessible Media*, 11(2), 33–49.
- [13] López, J., Fernández Torné, A., & Orero, P. (2022). Enhanced audio description: Towards immersive accessibility. *Journal of Visual Communication in Medicine*, 45(1), 24–35.
- [14] Natalie, R., Chang, R.-C., Sheshadri, S., Guo, A., & Hara, K. (2024). Audio description customization. *arXiv*.



- [15] Pavel, A., Reyes, G., & Bigham, J. P. (2020). Rescribe: Authoring and automatically editing audio descriptions. In *Proceedings of the 2020 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 1–13).
- [16] Perego, E. (2018). Into the language of museum audio descriptions: A corpus-based study. *Perspectives: Studies in Translation Theory and Practice*, 27(3), 333–349.
- [17] Remael, A., Orero, P., & Vercauteren, G. (2021). *Audio description: New perspectives illustrated*. Routledge.
- [18] Rosenberg, R. S., Kim, S., & Ahmed, M. (2024). Accessibility barriers in live audio chat applications for visually impaired users. *Journal of Human–Computer Interaction*, 40(2), 131–149.
- [19] Statista. (2024). Podcast listeners worldwide 2024.
- [20] United Nations Convention on the Rights of Persons with Disabilities (UN CRPD). (2006). Article 9: Accessibility.
- [21] W3C. (2023). Media accessibility user requirements. World Wide Web Consortium.
- [22] W3C. (2024). Accessible media guidelines: Audio description for podcasts. World Wide Web Consortium.
- [23] Wang, C., & Taelle, P. (2025). AI for accessible education: Personalized audio-based learning for blind students. *Journal of Inclusive Technology*, 17(1), 24–41.
- [24] Wolfswinkel, J. F., Furtmueller, E., & PH. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45–55.