

The Role of Voice-Activated Technologies in Shaping Podcast Engagement and Service Satisfaction: A Theoretical Synthesis Through the HMC Framework

Dr. Praveen Kumar Pandey (*)

Postdoctoral Researcher, Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia

Orcid Id - [0000-0002-9256-0093](https://orcid.org/0000-0002-9256-0093)

praveen.pandey2022@gmail.com

Prof. (Dr.) Sailesh Suryanarayan Iyer

Adjunct Research Faculty and Supervisor, Lincoln Global Post Doctoral Programme, Lincoln University College Malaysia, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia

Professor & Principal, Narnarayan Shastri Institute of Technology-Institute of Forensic Sciences and Cyber Security (Affiliated to NFSU, Gandhinagar) Ahmedabad

Orcid Id - [0000-0002-0619-9829](https://orcid.org/0000-0002-0619-9829)

drsaileshiyer@gmail.com

Corresponding Author (*) - Dr. Praveen Kumar Pandey, Postdoctoral Researcher, Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia & can be reached at

praveen.pandey2022@gmail.com

Abstract

Within about 10 years, two things took place. Podcasting suddenly got mainstream and taking dictation with a device became a smooth turn to digital content. Both the research literatures have hardly crossed paths. This paper attempts to take a step towards filling the gap with a conceptual model of the change that occurs when the initial signal for listening shifts from a tap to a spoken signal. This is where the Human-Machine Communication framework comes in handy as it gives an insight into why users tend to forget the machine's existence most of the time until it does something to fail them. There are two tech inputs, the first being voice added to the model and the second amount of platform technology intelligence. Both argue that they affect the listeners' expenditure out of that service, that is the usefulness of the service, the quality of the service and the interacting with the service. Those judgments in turn, lead to engagement and satisfaction. Technology readiness is a moderator that is parallel to the model, as it won't do the same thing for the same user. The big picture is related to SDG 9, and the issue of trustworthy conversational infrastructure.

Keywords: Voice-Activated Technology; Podcast Consumption; Human-Machine Communication; User Engagement; Technology Readiness.

1. Introduction

If you listen to their last podcasting show, ask a typical listener what they did to begin and a lot will reveal they spoke to a speaker. It sounds like a minor change of custom, but that is only on the surface. It probably is not. A majority of the smart speaker owners who use spoken commands say they use it to access podcasts, often while undertaking other tasks, whether cooking or cleaning or getting ready for an outing, NPR and Edison Research (2022) report. Listening is now when you're not doing anything else, the device that fills in these moments is one you converse with.

At the site of discovery, differences in the number are most apparent. If you own a cell phone, you find a few descriptions, pass on four items until you have arrived at the one item that you want. Approach an assistant and he or she will give you one suggestion if that; It's taken by most people. While Choice was not the end of the world, there is something else that unsartened it for you, and which alters what it means to have a choice in that case.

There is a ton of research on podcasting engagement. Have voice assistants as well. This obviously occurs seldom if ever these two appear in the same paper. Human machine Communication (HMC) is a good method to connect them, as it means users don't need to be aware that they're dealing with a machine all the time. Lee (2024) takes virtually the converse position, and states that people focus on the message and then forget about the source until the expectation is violated. It's an unusual example here for voice assistants. No one believes that Alexa is human. But many of the design elements suggest conversation – and the moment a title is misheard the machine ceases to be mere background, and becomes the subject of your interaction.

The following is a model of instead of a test on one. Voice interface quality and smart features integration are considered the technology inputs and agreed through three user judgments, culminating in engagement, and satisfaction, and mediated by technology readiness, the strength of the relationships. The overarching stake is of SDG 9. Voice platforms are no longer restricted ad hoc, but are now a part of the infrastructure that transports cultural content, and the credibility of the communication infrastructure now hinges on design decisions involving reliable and transparent platforms (Shin, 2025).

2. Review of Literature

2.1 Podcast listening, engagement and satisfaction

Podcast listeners are most certainly not like a typical media audience. Sessions last up to 2 hours, everybody focuses on the activity, and the affinity to a specific host may become as high as that towards a friend. This is because, as Chan-Olmsted and Wang (2022) explain, there are a variety of motivations to listen: for information, for entertainment, for listening when doing non-entertaining things (e.g. chores), and because no one else does. Being engaged by that work involves frequency, depth and the parasocial pull. All of it produces satisfaction and satisfaction is the reason why individuals revisit. The problem is that nearly all of these studies feature some individual with a cell phone in hand. Remove the screen and the technology leaves the "background. starts to make choices.

2.2 How does the device change upon listening back

Today, smart speakers live within the places they hear their owners use them, and not in a specific listening spot (NPR & Edison Research, 2022). A look at where voice-based recommendation is going may highlight what's cause for concern. Those that hear suggestions think of less options and accept the initial ones much more easily than those who are reading a list. Explore less, be convenient more. For this model that provides an action in two-way. The use of easier access should lead to increased perceived usefulness. Provided authenticity as well as a less strong feeling of choice may subtly erode the sense of satisfaction of the service. Also introducing some type of transparency and variety is not a nicety design. It is one of the elements of the infrastructure's trustworthiness (Shin, 2025).

2.3 Voice interface quality

The 'voice interface quality' here refers to how language is presented when speaking. Is it possible to catch the words (including accented words and words over a running tap). For how long does it take to get the answer? Is there a "voice problem" (grainy voice). Looking into the VUI literature, one thing is clear, as Deshmukh and Chalmers (2024) report – accuracy and low effort have great impacts on customers building trust, while repeated failure will have them return to the app.

Perceived usefulness" as defined by Davis (1989) is the belief that the system enhances your performance and "interface reliability" is directly contributing to it. If a speaker knows what episode you're listening to, and can catch up with you there, you don't have to play catch-up and hands-free listening is just possible. The reliability component also translates to the concept of service quality because reliability and responsiveness are measures of the any technology mediated service (Parasuraman et al., 1988). Then the softer, meaning did the exchange feel easy/hard.

2.4 Smart features integration

Smart features integration – amount of intelligence behind the voice. Personalized suggestions, based on your browsing history. Any suggestions which vary according to the time of day. Playback you can enjoy throughout the house, from the kitchen to the car and that can be integrated into anything that runs in the house. When it works, this is the search effort relinquished and the service feels useable - again as usefulness. Consistency relates to reliability and assurance, two components of service quality (Parasuraman et al., 1988) and is a sign of competence. The use of one of these vaguely worded requests results in an intelligently worded reply from which the user infers that the reply has been understood. Rawool et al. (2025) are the ones that just make this impression – they relate it to the trust that users have in voice assistants, and whether they continue to use them.

2.5 Human-Machine Communication

HMC's main thesis is that source's awareness is not continuous. This audience does not come to watch the machine but for what it says; expectations need to be broken until then. (Lee, 2024). Once it takes a stance, then an existing belief about machines takes an interpretive stance. Basic voice listening is a fine instance. While playing the episode no one is thinking about the assistant. Judgement about the episode isn't really about the episode in the first second after the mishearing and the mishearing 'changed' the name!

2.6 Technology readiness

Readiness is not an ability, it's an attitude. That's optimism and innovativeness tugging in the opposite direction of discomfort and insecurity, says Parasuraman and Colby (2015). Advanced features are trying to reach out to ready users – and shrug off the occasional fail. For less ready users this failure is considered confirmation of their ability to have been misguided. However, the reason for the inclusion of readiness here as a moderator rather than another predictor is because of the fact that readiness is embedded in the schools' neurological activities. If you are someone who already is keen on using a product, the better the interface is, the more valuable it is.

2.7 Judgements to engagement and satisfaction

The strong, relatively stable results relate to uses of perceived usefulness to predict continued use (Davis, 1989) and services to behaviour and to satisfaction (Parasuraman et al., 1988). What those two lack is the quality of interactive experience – was it fun or not? Communications are

easy and generate longer listening times to warm up people to the platform (Chan-Olmsted & Wang, 2022). Then engagement again goes back to satisfaction as they are likely to give if they are rating the service (Oliver, 1999).

3. Proposed Hypotheses

H1: Voice interface quality positively influences perceived usefulness, perceived service quality and interactive experience quality.

H2: Smart features integration positively influences perceived usefulness, perceived service quality and interactive experience quality.

H3: Perceived usefulness, perceived service quality and interactive experience quality positively influence user engagement.

H4: Perceived usefulness, perceived service quality and interactive experience quality positively influence service satisfaction.

H5: User engagement positively influences service satisfaction.

H6: Technology readiness positively moderates the relationships between voice interface quality and the three user evaluations.

H7: Technology readiness positively moderates the relationships between smart features integration and the three user evaluations.

4. Conclusion and Future Directions

The model presented here is a conceptual model, and is intended for testing, not admiration. What it really says is that 'voice quality of the interface' and 'inclusion of smart features' is not the same thing, they're not a single technology variable, anyways, and really, readiness is what matters and determines how much the voice quality of interface or the integration of smart features matters to them.

Testing of it would require a survey of smart speakers users, as well as SEM model containing the interaction term of readiness. Comparing between users who use voice primarily (as opposed to using an app) would help isolate the interface from tons of other factors. The weakest statement of the argument is the statement about the exploration claim. It took some research by the podcaster, and it was not seen in podcast listeners' behaviour, but data on the extent of voice first listeners would be interesting to collect. It's a matter that transcends "user satisfaction" if it is indeed a concern. Quite seamless and contracting system is not the type of infrastructure SDG 9 should be.

References

1. Chan-Olmsted, S., & Wang, R. (2022). Understanding podcast users: Consumption motives and behaviors. *New Media & Society*, 24(3), 684–704. <https://doi.org/10.1177/1461444820963776>
2. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
3. Deshmukh, A. M., & Chalmeta, R. (2024). User experience and usability of voice user interfaces: A systematic literature review. *Information*, 15(9), Article 579. <https://doi.org/10.3390/info15090579>
4. Lee, E. J. (2024). Minding the source: Toward an integrative theory of human–machine communication. *Human Communication Research*, 50(2), 184–193.

5. NPR, & Edison Research. (2022). *The smart audio report 2022*. Edison Research. <https://www.edisonresearch.com/smart-audio-report-2022-from-npr-and-edison-research/>
6. Oliver, R. L. (1999). Whence consumer loyalty? *Journal of Marketing*, 63(4_suppl1), 33–44. <https://doi.org/10.1177/00222429990634s105>
7. Parasuraman, A., & Colby, C. L. (2015). An updated and streamlined technology readiness index: TRI 2.0. *Journal of Service Research*, 18(1), 59–74. <https://doi.org/10.1177/1094670514539730>
8. Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.
9. Rawool, V., Foroudi, P., & Palazzo, M. (2025). AI-powered voice assistants: Developing a framework for building consumer trust and fostering brand loyalty. *Electronic Commerce Research*, 25(6), 4471–4503.
10. Shin, Y. (2025). Toward human-centered artificial intelligence for users' digital well-being: Systematic review, synthesis, and future directions. *JMIR Human Factors*, 12(1), Article e69533. <https://doi.org/10.2196/69533>