

Explore the Dynamic Effects of Audio Voice on Podcast Listeners' Positive Emotional Responses

Jun Ji¹, Kotchaphan Youngmee^{2*}, Khachakrit Liamthaisong³

¹ Department of New Media, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand. Email: 51431046@qq.com

^{2*} Department of New Media, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand. Email: kotchaphan@msu.ac.th

³ Department of New Media, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand Email: khachakrit.l@msu.ac.th

Corresponding author: Kotchaphan Youngmee, kotchaphan@msu.ac.th

Cite this paper as: Jun Ji, Kotchaphan Youngmee, Khachakrit Liamthaisong (2024) Explore the Dynamic Effects of Audio Voice on Podcast Listeners' Positive Emotional Responses. *Frontiers in Health Informatics*, 13 (3), 5062-5079

ABSTRACT

In this study, the researchers directed their attention towards the phenomenon of sound. Specifically, they sought to examine the emotional responses of the participants by analyzing electrodermal responses (EDRs). The study involved the comparison of two distinct sound conditions. Initially, it was observed that there was no discernible variation in the Electrodermal Response (EDR) pattern across different conditions. However, upon careful examination of the peak EDR outputs observed in various simulations, it became evident that the emotional content had a discernible impact on the spatial conditions. Furthermore, it was particularly noteworthy that the interaction between sound and the aforementioned spatial conditions played a significant role. It was observed that no distinct individual stimulators could be identified, although there were notable trends and variations in Emotional Dysregulation Rating (EDR) under specific circumstances. In addition, it is worth noting that the participants in the study recognized the most prevalent instances of positive emotions across all experimental conditions. Furthermore, the utilization of a combination of arousal sound design and 3D sound mix elicited the greatest number of emotional responses. Conversely, the omission of data regarding participants' educational level, cultural background, and gender prevents a comprehensive exploration of these factors, thereby limiting the potential for generating more nuanced insights.

KEYWORDS

Emotional responses, electro-dermal responses (EDR), sound conditions, stimuli, positive emotions

Nomenclature

$K\Omega$ Kiloohms

μ Mean

x Individual values

1 Introduction

1.1 Background and Overview of Podcasting

Podcasting is the one media niche that has witnessed significant and rapid development right within a period of twenty years, and this has changed it from being a method practiced by a few enthusiasts, into a worldwide phenomenon. In 2004, the throng of Google search results for the term 'podcast' amounted to 6000. But in 2005, that number suddenly grew by 60 million or so [1], and now, the number of 1.9 billion has been reached. The blossom of listening hours echoed the growth recorded in the number of listeners and content created on a global sphere respectively [2]. The future of podcasting may be a paradigm of evolution, stability or even contracting, but then it invariably allows for podcasting being flexible and adaptive to new trends and later expectations.

Moreover, the development of tools and technology in sound will be critical in bringing about a new era of greater involvement of the audience and professional output of the content [1]. Beniamini studied the probability of at least one podcast listening for a person in the US [3]. The frequency in terms of lifetime habit was 0.55. Yet, though podcasting is a widely accepted medium, many still fail to closely associate the term 'podcast' with a specific type of audio content.

However, the mixed feelings about conceptualizing the podcast structure are further aggravated due to the intermittent advances in technology and the possibility of updating the established structure thereby upgrading the whole podcast ecosystem [4]. This poses a crucial inquiry: how can we develop podcast production and usage as the medium itself or what it symbolizes stays unchanged? For us to handle this hypothesis, it would be necessary to make an indication that is precise and particular on the term 'podcast'. The definition and limitations of podcasting can be thoroughly explored only if we observe and measure the underlying tensions that permeate through the very nature of the medium. This study is tailored to cover topics that are yet under-studied in the field of podcasting research and therefore used to find the direction of the platform and how it is continuously evolving.

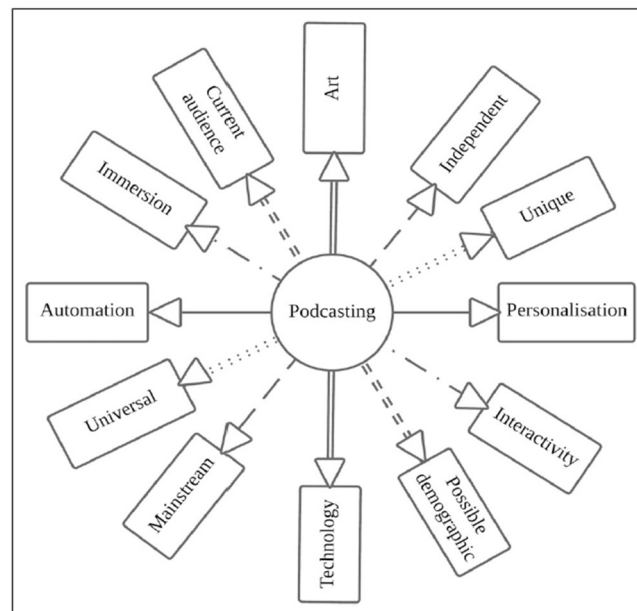


Figure 1

Using six pairs of concepts important to the medium, the six-tension framework for podcasting innovation revolves around the concept of podcasting. It is postulated that to maintain the essence of podcasting, it is necessary to maintain the equilibrium within the pairs during any form of innovation. [1]

1.2 Importance of Audio Voice Characteristics in Podcasting

Audio-related qualities of voice are an essential component that not only attain the right sound and as a result save the listener's attention for the content but also can turn an audible one. The spoken quality is the way words are tipped or spoken which is made up of difficult vocalizations and other pronunciation features which are the definers of the perception of sounds greatly. Humanizing: The power of persuasion is very centered, that is shuffle-down/ its focus is on capturing the listeners' concentration level and thus, has an immediate influence. This message connects deeper, hence the audience feels unequivocal affection towards it. Similarly one of the consequences of leveraging monotonous and irritating voice is the tendency to lack of sustaining the focus of the audience and its forgetfulness of the core message.

Lastly, the clear fact that an individual's tone which usually accompanies a high power of voice, can greatly influence opinion as it may make people believe and also remember information that is brought across that way.

A voice that quite skillfully and enthusiastically reflects the keys of the emotions can powerfully impinge the emotions of the audience, condition the reflection, and provoke the audience to critically think. This option allows for the check of the highest precision of genuine message content. Communication of tone or intricacies of matters adds sparkle and makes things a lot more vividly informative. However, a speaker whose topic is dull or without excitement will probably mislead the audience and they may thus be confused enough, enough that there would exist lesser connection and lesser involvement.

The aural clause therefore contributes to establishing the importance of language learning and it shows a key aspect of communication study. Through podcasting, students manage to bring out little details impress, enhance compliance, and formulate their way of vocalizing. Using self-examination of their voice as the communication means as well as a leitmotif of their communication with the community, students can explore how linguistic aspects of a speech appear when being delivered and how they are taken on. Through this reflective journey, which involves the analysis of paralanguage and chronemics, the students will be capable of deciphering nonverbal communication's importance as a delivery medium of message.

1.3 Research Gap and Purpose of the Study

These gaps in this field of work arise because there is not enough information and validations justifying their research on the influence of audio voice characteristics and emotional reactions of audiences in podcasting. While podcasting is the most widely used method to transfer pictures and quite a lot very science in society at the moment, all matters, the specifications of the sort of surfacing modifications could to the feelings of the audiences are not pertinent. Furthermore, whereas there exists some research on how the voice with storytelling and material skills influences the sentiment of the audience of podcasts, still, the particular exact voice impact on the audience has not been investigated. Language is a great tool that media use in direct communication with the audience to convey the message. What should be the main focus of podcasters is that they take into account the role of the different voices of several hosts, with various genders, nationalities, and backgrounds, on the emotional responses of their listeners.

Moreover, the current body of "evidence" concerning audience engagement indicators like download figures, subscriber numbers, and listener feedback explores exclusively the surface-level engagement mechanisms, glossing over more profound psychological mechanisms underlying emotional connection. Hence, the need for investigations using empirical approaches that are based on a properistic approach, such as experimental designs and qualitative analyses, is eminent to illuminate the link between sonic voice attributes and the feelings of their targets in podcasting.

1.3.1 Research Objectives

- i. To examine how different podcast hosts' voices and their genders influence positive emotional responses from listeners.
- ii. To develop an audio voice model to reflect the audience's Positive emotion for podcasting
- iii. To compare the positive emotions of the audience before and after listening to a podcast.
- iv. To study the relationship between voice models reflecting positive emotions and the gender characteristics of podcasters.

1.3.2 Research Questions and Hypotheses

The following research questions will be addressed within this thesis:

1. What is the impact of various audio voice characteristics on eliciting good emotional responses from podcast program audiences?
2. What is the correlation between the gender of podcast hosts and the occurrence of favorable emotional responses from listeners?
3. How much do listeners' positive emotional responses differ before and after listening to a podcast?
4. What is the relationship between the audio voice model, which is created to mirror the positive emotions of the listener and the gender characteristics of podcast hosts?

1.3.3 Hypothesis:

1. Null Hypothesis (H0): ASD has no recognized statistical bias in listening to favorable emotional reactions among the audience toward the quality of audio voices.
Hypothesis 1: By planting distinctive music voice qualities, eliciting good emotional reactions from the listeners means a lot.
2. Null Hypothesis (H0): Whether podcast broadcasters are male or female, such does not mean there will be a big difference in listeners' positive emotional responses.
H1: It is women presenters that are associated with a significantly increased variance of positive emotional responses.
3. Null Hypothesis (H0): Even though the effect of podcasting on the amusement of listeners' emotions is not statistically significant among listeners before and after the podcast.
H1: There is a marked contrast between positive emotions and reactions during the time you listen to a podcast.
4. H0: We can conclude that there is no significant correlation between the audio voice-over mode which reflects the positive emotions of the targeted audience and the specifics of the gender features of podcast presenters.
H1: The gender characteristics of podcast hosts are influenced considerably by the emotional portrayal of the audio voice model, stirring a pleasant sentiment within listeners.

1.4 Research Scope

1.4.1 Population:

The population under study is the community of podcast program listeners. This involves users who regularly tune in to podcasts of different themes and formats either on different web platforms or channels that are easily accessible.

1.4.2 Sample Group:

The recourse of a sample group of 50 people will be applied through volunteer sampling. The sample would comprise 25 male and 25 female colleagues and thus, would be demographically and tastefully diverse.

1.4.3 Research Tools:

The study will utilize a variety of research tools to collect data and analyze the relationship between audio voice characteristics, gender, and audience emotional responses:

- 1.4.3.1 *Questionnaire*: The survey will be structured to catch demographic information as well as evaluate listeners' emotions prompted by podcast content.
- 1.4.3.2 *Interviews*: Interviews will take place remotely or in person depending on a researcher's requirements and it will help to get a better understanding of listeners' perceptions of audio voice characters and the type of gender of podcast hosts.
- 1.4.3.3 *Quality Format Assessment Survey*: Since the survey is mostly about podcast format, it will be created to measure the coverage including audio production, content relevance, and the level of satisfaction.
- 1.4.3.4 *Positive Emotion Measurement Form*: The instrument form will be applied to measure participants' emotional responses toward the podcast episodes impressing them in the pre and post-listening.

1.4.4 Top 10 Most Popular Positive Podcasts:

The study will analyze the top 10 most popular positive podcasts, organized by different genders of hosts:

- 1) Top 5 podcasts hosted by males
- 2) Top 5 podcasts hosted by females

These podcasts will serve as the primary content for assessing the influence of audio voice characteristics and gender on audience emotional responses.

1.4.5 Variables:

The study will examine the following variables:

1) Independent Variables:

- 1.1) Audio voice model to reflect the audience's Positive emotion for podcasting
- 1.2) Gender: Female and Male Podcasters

2) Dependent Variables:

- 2.1) Positive Emotional Responses Scores of Podcast Audience
- 2.2) Relationship between an audio voice model to reflect audience Positive emotion and the gender characteristics of podcasters

Literature Review

Definition and Conceptualization of Audio Voice Characteristics

Talking tone possesses a particular sonic features usually described as speechprint that provide the source of information that is heard exclusively by the target audience. Context qualities of voice are reaching out by expressions which include pitch, tempo, tone, and various characters of vowels as well as articulations comprise these traits. Soundtrack' is a new performer that incorporates the vocal basing the whole performance on several factors. The voice tone is a process of dealing with many aspects of the voice: such as the rhyme in the verse, the rhythm of the verse, the voice intonation, its friendliness or commanding nature, and its warmth. Here, branding is used as a link that helps in arousing the customer's desire.

This brochure will use the word 'speed' to explain how fast "the speaker utters the thoughts" or what is expressed in that context which I will be referring to as pace. The tempo of slow is used to convey important information slowly or gracefully, which is the opposite of acting drama in compliance with a faster tempo. It can be said that in proclamations or role-plays differentiation of modulations will be necessary as it will deepen acoustical perceptions (proclamations: volume, intonation, pace, and voice) with the help of rhythm and temper. In particular, Neil worked to improve the technique of reading, and the ease of understanding, the listeners' orientation, same as.

Even more, the devotedness to this audio world dictates the direction in which podcasters are going to make their fortune and the spoken quality could be very useful as well as determinant though if not done in good ways nothing positive will come out of the business undertaking because that is how the audio quality adds flavor to any data that does not bore the listener. They accomplish this with more than dissemination but also theories of action as well as practice on the part of viewers, through adapting personal level is being true, and triggering viewers feeling.

It is logical to say that the listener's consideration is the source of all this. The experience contributes highly as a tool for fostering the truth and effect of the communication. This consequently provides the podcast hosts an opportunity to make public facts or an elaborate statement thereby gaining their audience and holding them to the listening process. The podcaster also creates genuine connections with the listeners of the show. Besides, the podcaster is made to always improvise and keep their listeners excited. Not to mention that they can add an extra touch of details to make the story more intense and their audience more taken in even more.

Psychological Foundations of Emotions in Audio Perception

Emotional interpretation in audio relies on a wide range of psychological mechanisms that impact how people take in and process auditory information. It is crucial to develop a comprehensive understanding of how the voice traits in audio impact the emotional responses of podcast listeners. To achieve this it is important to take note of the those key points. When hearing sounds we tend to feel certain emotions that they evoke. The response to vocal stimuli and the ability to emit auditory effects among people is topic that would contribute towards the fulfillment of this research's 'objectives. The two illustrate the manner in which audio information may evoke emotions by tapping into ones subconscious memories. A persons comprehension of sound perception is greatly impacted by the notion of psychosomatic emotion contagion. People usually absorb and radiate the feelings of their surroundings, especially those that are transmitted through subtle nuances or voice

tone. Psychologists denote this phenomenon as emotional contagion, where listeners find themselves emotionally linked to the speaker's expressions subconsciously.

Additionally, a tool that would also be valuable in the comprehension of how listeners perceive and evaluate the quality and value of podcast voices is the appraisal theory of emotions. This theory illustrates how individuals assess auditory inputs consequently associating emotional values to them, whether negative or positive, similar to other surrounding signals. There is a spectrum of cognitive filters at play in the course of this process, which tend to filter out given information whilst intensifying emotional reactions.

The cultural context of emotional responses to audio data is also an important aspect to look into. Looking into speech attributes a range of factors like societal standards Personal differences, and cultural norms have a potentially great impact on how the listeners perceive it. When considering podcast listeners, their understanding of the content and connection with the podcaster may be influenced by societal norms regarding communication and emotional expression. There are a lot of factors, including evolutionary, social, cognitive, and cultural ones, which contribute to how people interpret and respond to auditory stimuli.

Studies on the Influence of Audio Voice Characteristics on Emotional Responses

Podcast journalism is an exceptional tool where audio is the backbone of the making telling of engaging stories. In an investigation conducted in 2019 involving the most rewarding podcasts in the UK and Australia, it is recommended that journalists strive to create an emotive connection with listeners as a necessary but sufficient paraphernalia for any podcast to become significant [5]. Podcast journalism deviates from the traditional approach of objective reports by championing the humanity of others and person-to-person stories, creating a context for both listener and speaker to open up and reflect.

On voice persuasion, studies show that some vocal states (such as low intonations, increased pitch, as well as pauses) have a considerable effect on emotions and perception in specific situations, but it is not directly associated with podcasts [6]. Being able to achieve the results by varying the loudness, pitch, and tone of the voice is a prerequisite for a better communication rebuilding and increasing of your arguments' persuasiveness. The role of audio voice characteristics in developing rapport and engagement of audiences is portrayed in the fact that magazines that use first-person narratives are more likely to listen to them [7]. The voice of the storyteller has an irreplaceable part in determining how the listeners perceive the content and what their experiences of it will be like [8]. It is also possible to use this type of interaction to build personalized stories or at some point to start having a parasocial interaction.

Podcasting as a Medium for Emotional Engagement

Besides, a significant proportion of the results obtained from the research related to this idea have been proof that the helping of this type of media in growing the listeners' sentimental feelings is an essential matter. Scientist experts worked on their research called the typical Australian and British podcast [5], to ensure that it considered that the podcast type was effective when it comes to private situations and a real live telling of the conservation. Spoken language and the way journalists use flashbacks and other sound work tools create good relationships that in turn keep the audience engaged and excited and thus compel the feeling of belonging and companionship.

Advertising via podcast is not the only positive feature of this setting; it is also a kind of communication from the company to its audience. Radio format usually is characterized by a loyalty level much higher than satisfied with television viewership and even memory. Studies proved that podcast listeners tend to have a longer recall and more intense emotionality when listening to their favorite brands being mentioned [9]. The Economist, The New York Times, and The Guardian are just some of the podcast publishers grouped in this category, the others aiming at the same goal; to gain new listeners and reinforce the engagement of existing ones. While listening to podcasts, one experiences, somehow, a parallel invitation toward both the known and the unknown. Together they create a strong storytelling, which then turns into a story that unfolds in a manner that enables the listeners to personally connect with it. The podcast is a perfect medium for brands, personalities, casters, or other storytellers to gain their followers' notice--via the listeners' hearts.

Theoretical Frameworks

This paper will explain how to evaluate the effect of audio voice attributes on emotions in podcasting,

through which the subject can be interpreted with the use of a wide range of theoretical models. Researchers can use these frameworks as tools for data processing to help build connections between the affective impact of podcast content, sound variables, and voice characteristics.

- 1) **Social Presence Theory:** Social presence theory suggests that the operations that involve the communication of social cues and the interactions that are enabled via media differ. Voice tone, pitch, and modulation are the attributes that form the podcaster's social presence the article shows [10].
- 2) **Elaboration Likelihood Model (ELM):** The ELM is a theory that provides insight into how a person processes any message of persuasion. The Elaboration Likelihood Model posits that individuals can be persuaded by persuasive messages through two distinct routes: the main respective of the type either central way and the peripheral route [11]. This kind of research can uncover certain characteristics of podcast listeners' impressions and emotions as far as the ELM framework is established.
- 3) **Parasocial Interaction Theory:** The Parasocial Interaction Theory is Daniel M. Giles's explanation of the formation of one-sided relationships between individuals and media personalities and how individuals interact with people who consume the media. Listening to a podcast can eventually lead to building a parasocial relationship as people may start relating to their host. It could be the feeling of family and understanding or their similarities with them.
- 4) **Emotional Contagion Theory:** The emotional Contagion hypothesis holds up the idea that humans can learn or copy the emotions of others by following their body language, facial expressions, and the slightest changes in the intonation of their voices [12]. As podcast listeners become more spiritually connected with the host because of her ability to express what she means through her voice they may feel very emotional about what they learn.

Research Methodology

Research Design

The research technique employs a mixed-methods approach to investigate the impact of audio voice characteristics on the audience's positive emotions in the field of podcasting. This is achieved by conducting four experimental components. The primary goals of this study chapter are to investigate and comprehend the varying effects that the quantity and kind of podcast hosts have on emotional audience responses. Regarding the research, it is most likely that the subjects are exposed to selected portions of podcasts when the hosts demonstrate their vocal characteristics and gender.

The second experiment seeks to predict the lavishing emotions that listeners feel while they comprehend the material offered on the podcast. This procedure aims to leverage machine learning approaches to highlight the assessment of audio traits, their relationships, and the emotional responses among the listeners. The chief objective is to establish a predictive framework that establishes relationship between audio traits and the effective emotional connotation within the realm of podcasting. Furthermore, the data is keenly assessed to establish the degree to which individuals are affected by their own emotions. The consequent experiment, looks into the link between the voice model scenarios in experiment. These scenarios encompass both female and male bias and their effect on podcasters. By evaluating the effect of a hostess on the audiences' perception of gender, researchers scrutinize the prospective impact of preference or bias theories on audience emotional reflexes.

Participant Selection and Recruitment

The sampling method was utilized in the recruitment of the participants, with keen focus on convenience. This study focused on a spectrum of persons from diverse episodes, leaning towards leisure activities, well-being

and entertainment disregarding drama since it may not be as appealing as the rest. The aim was to come up with an assortment of 50 individuals, with an equal fraction of the two genders. Hence, the findings within this study will avail a precise representation of the diverse groups present amid the separate populace.

Stimulus Materials

In the course of this study, a group of stimuli was crafted and the team handpicked the top ten podcasts. The study assessed the languages used in the selected podcasts, not forgetting gender disparities amid the hosts. It became possible to exploit various concerns only after incorporating a diverse assortment of podcast presenters and genre, rather than only looking into one or a few. The final experimental stimuli were meticulously chosen to showcase a wide range of emotional responses, effectively eliciting them from the participants.

The stimulus materials may have been selected with the intention of eliciting a variety of emotional responses from the spectators, possibly due to an underlying deeper significance. This could involve selecting podcast episodes or snippets specifically advised for their cognitive impact in enhancing positive emotions, such as enjoyment, motivation, and amusement. Sufficient emphasis may have been placed on the style, speed, and subject matter of the chosen materials and study goals, and their impact on the desired outcomes.

Data Collection Procedures

Data acquisition encompassed a diverse range of activities, including the measurement, evaluation, testing, and collection of information from participants both before and after their exposure to the podcast stimuli. The tools utilized were real-time measures of participants' emotional reflexes and perspectives, quality format assessment surveys, interviews and questionnaires. This statistical approach avails a holistic collection of information to be used to derive emotional connotation.

The questionnaire and the survey were utilized in the collection of quantitative data, where the two were carefully molded to prompt organized answers from the individuals concerning their emotional stance, the quality of the media, and their holistic podcasting experience. This encompassed Likert scales, making use of numerical values to gauge the participants' emotions ranging from degree of agreement with the media, satisfaction, and happiness. On top of the quantitative data, qualitative insights were also recorded by interviewing the participants. These interviews allowed the interviewees the chance to express their, decisions, feelings, and view clearly. Consequently, the main aim was to gain a vivid psychological acknowledgement of the interviewees. This includes their experiences and the aspects that led to their recorded reactions to the media.

Data Analysis Techniques

The team utilized a meticulously curated methodology to avail quantitative information concerning the extent audio voice qualities entice positive emotions among the audience within the context of podcasting. The efforts incorporate both quantitative and qualitative methods to develop a holistic understanding of vast levels of emotional reflexes impacted by vocal traits. This paper utilized descriptive statistic as one of the analytical tools seeking to avail an accurate overview of the emotions of the audience and their podcast quality ranking. This method aims at providing statistical information concerning the central tendency and distribution of the data. Correspondingly, the inferential statistics was used to evaluate the link between variables, mainly leaning towards the effect of voice qualities and gender of podcast hosts on the positive feelings of the audience. This assessment will show statistical evidence of the vibrancy of these contributions.

In regards to the qualitative method, a thematic analysis will be utilized to identify key patterns and subjects present in the text of the qualitative data obtained from questionnaire interviews or open-ended survey questions. This approach facilitates the identification of subtle nuances in the emotional state of listeners and their

perception of what they are contemplating. It also enhances the depth of understanding of the underlying mechanisms in the method of analysis. The report will incorporate both quantitative and qualitative analysis to ensure methodological consistency and logical coherence.

Results

Descriptive Statistics

In this section, descriptive statistics are derived from the data obtained in the study. The analysis takes into account the impact of voice characteristics on emotional responses in podcasting. An analysis was conducted on pre- and post-emotion reaction ratings, revealing that the average pre-listening response was 4.2, while the average post-listening response was 4.8. The average of the emotional response range before hearing was 0.6, while the emotional response range after listening was 0.4. The impact of listening on emotion was apparent in the diverse range of pre-listening emotional reaction ratings, which ranged from 3 to 5, whereas post-listening emotional response ratings ranged from 4 to 5 on the scale. Furthermore, the assessors' vocal qualities were assessed based on mean evaluations for tone, pitch, speed, and modulation. The statistics can be summarized as follows: the tone is rated at 4.6, the pitch at 4.3, the rate of speech at 4.5, and the modulation at 4.7. The acoustic properties had standard deviations of 0.5, 0.4, 0.6, and 0.3, respectively. Example sentence: The job market has been significantly impacted by automation in recent decades. Over the past few decades, automation has had a significant influence on the job market. Regarding the gender dynamics among the podcasters, the data revealed that 25 episodes were produced by males, while an equal number of 25 episodes were produced by females. The outcome led to a dispersal pattern that could then be represented by either a histogram, which uses pictures, or a bar chart, which uses bars, to depict the frequency distribution. Furthermore, the study examined the correlation between the ratings of the audio voice model and the gender of the podcast hosts. Summary statistics were then contrasted by analyzing episodes hosted by male versus female. The descriptive statistics serve as the basis for further analysis and explanation in the subsequent stages of the research.

The following formulas were used to establish the results within this section:

Mean: $Mean = (Sum\ of\ all\ values) / (Number\ of\ values)$

Range: $Range = Maximum\ value - Minimum\ value$

Standard Deviation: $Standard\ Deviation = \sqrt{\Sigma(x - \mu)^2 / N}$

In this case:

x = each individual value

μ = mean of the values

N = number of values

Analysis of the Impact of Audio Voice Characteristics on Audience Positive Emotions

The measurement of the electrodermal response (EDR) taken for evaluation of the phasic activity was found to be closely linked to the emotional state of each participant, as indicated by the electrodermal resistance measured in kilohms ($K\Omega$). Before the studio effect was introduced, all participants underwent a conditioning stimulus to familiarize them with the listening conditions. This allowed for the establishment of individual baselines for the EDR reactions. Baseline records unique to each participant were subtracted from the dataset of measurements. First, an ANOVA was conducted on the average EDR value throughout the stimulus period,

with sound type being considered as a factor that varies between subjects. However, the findings did not reach statistical significance ($p > 0.05$). Additionally, the results of the Mann-Whitney Test revealed a significant difference in terms of EDR between the groups ($p > 0.05$).

Peak EDR \pm SE values [k Ω] for each educational level (4th EP and 1st ESO) and sound condition (Neutral stereo, Neutral 3D, Emotional stereo, and Emotional 3D) for the different events (stimuli).

	Stimuli version	Condition A	Condition B	Condition C	Condition D
4 th EP	Meteor fall	0.912 (0.92)	1.027 (0.88)	4.115 (0.88)	0.78 (0.85)
	Dressing as scientist	3.176 (1.03)	3.394 (0.98)	1.794 (0.98)	4.19 (0.95)
	Tic-Tac	0.826 (0.90)	2.23 (0.86)	3.4 (0.86)	0.793 (0.83)
	Chronometer	1.778 (0.97)	1.628 (0.92)	3.543 (0.92)	1.726 (0.89)
	Sad day for light-speed	1.939 (0.99)	3.008 (0.94)	3.356 (0.94)	3.195 (0.91)
	Down effect	0.84 (0.83)	0.607 (0.79)	2.796 (0.79)	0.731 (0.77)
	Skate-down	1.707 (1.1)	1.859 (1.04)	2.293 (1.04)	2.846 (1.01)
	MEAN 4th EP (SD)	1.638 (0.79)	1.836 (0.99)	2.762 (1.08)	2.001 (1.29)
1 st ESO	Meteor fall	1.332 (0.90)	1.54 (0.91)	0.859 (0.89)	0.956 (0.87)
	Dressing as scientist	2.952 (0.99)	3.531 (1.01)	3.046 (0.99)	2.498 (0.96)
	Tic Tac	1.317 (0.87)	2.389 (0.89)	0.878 (0.87)	1.887 (0.84)
	Chronometer	2.602 (0.93)	1.816 (0.95)	2.578 (0.93)	2.348 (0.91)
	Sad day for light-speed	2.107 (0.96)	3.317 (0.97)	3.158 (0.96)	3.196 (0.93)
	Down effect	0.843 (0.80)	1.186 (0.82)	1.752 (0.80)	0.896 (0.78)
	Skate-down	2.231 (1.06)	4.196 (1.08)	2.639 (1.06)	5.121 (1.03)
	MEAN 1st ESO (SD)	1.801 (0.78)	2.378 (1.18)	2.025 (0.95)	2.224 (1.45)
	MEAN GLOBAL (SD)	1.72 (0.76)	2.11 (1.09)	2.39 (1.06)	2.11 (1.33)

Mean values (SD) for 4th EP and 1st ESO are marked in bold font.

Table 1 Educational Level

For this experiment, the EDR value was determined as the highest EDR observed during the duration of the stimulation. The table above, Table 1, presents the peak values of EDR for each grade (fourth EP and first ESO) along with their corresponding sound conditions (neutral stereo, neutral 3D, emotional stereo, and emotional 3D). Please make an appropriate reference to this table. The PEDR values were chosen as dependent variables for the Multivariate Analysis of Variance (MANOVA), considering the course (4th EP and 1st ESO class), sound emotional condition (neutral and emotional), and sound spatiality condition (stereo and 3D) as the between-subjects factors. The results indicated a slight inclination that did not reach statistical significance at $p = 0.07$ due to the interplay between the emotional sound condition and the sound spatial order condition.

Further analysis of EDR revealed that the peak value increase for neutral conditions was lower in the 3D context compared to stereo. However, the difference between stereo and 3D conditions was somewhat similar for emotional contexts. Our bivariate analyses revealed significant effects for some conditions, although not all. We did find certain trends for the sound spatiality test with the "pull around" stimulus. Specifically, the EDRs were generally larger for the 3D version, with a p-value of 0.085. Furthermore, the "tic-tac" sound demonstrated a potential connection between the emotional aspect of sound and its spatial characteristics, approaching a level of statistical significance ($p = 0.097$).

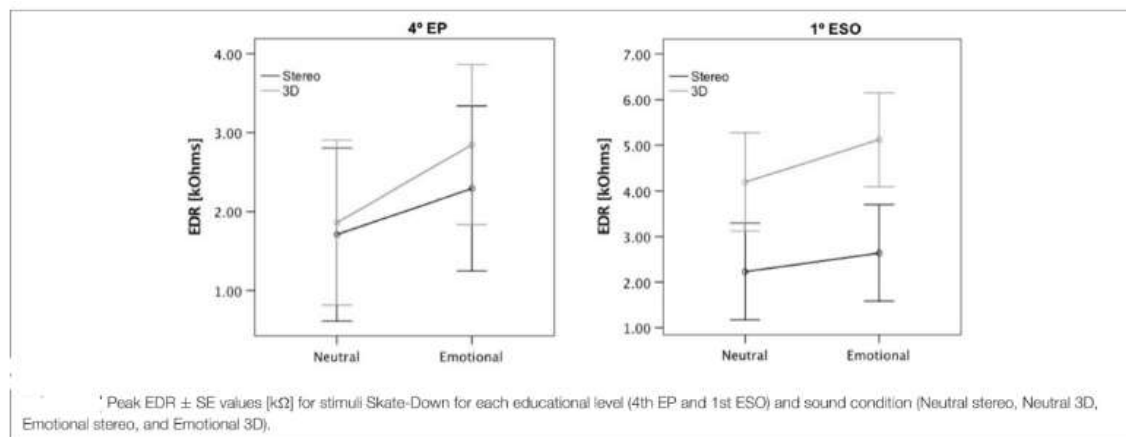


Figure 2 Stimuli Skate-down

Further analysis of EDR revealed that the peak value increase for neutral conditions was lower in the 3D context compared to stereo. However, the difference between stereo and 3D conditions was somewhat similar for emotional contexts. Our bivariate analyses revealed significant effects for some conditions, although not all. We did find certain trends for the sound spatiality test with the "pull around" stimulus. Specifically, the EDRs were generally larger for the 3D version, with a p-value of 0.085. Furthermore, the "tic-tac" sound demonstrated a potential connection between the emotional aspect of sound and its spatial characteristics, approaching a level of statistical significance ($p = 0.097$).

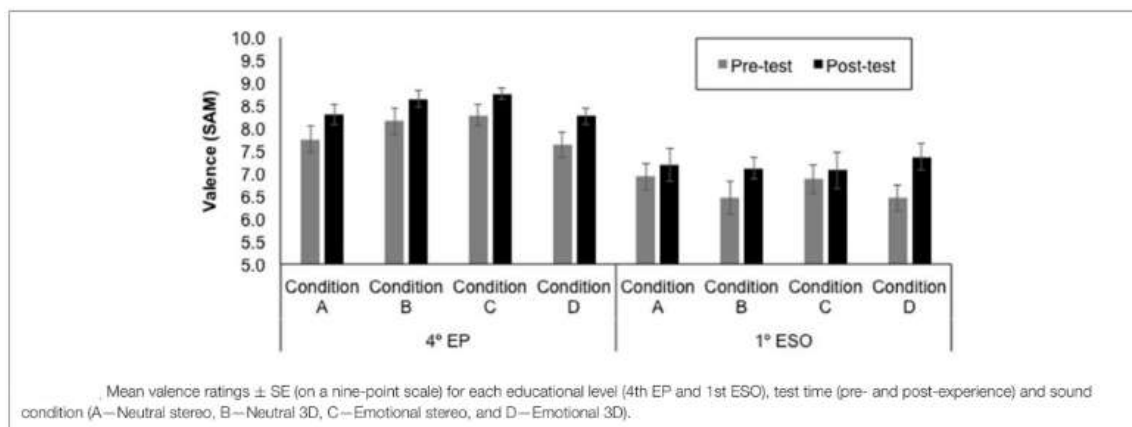


Figure 3 Mean Valence

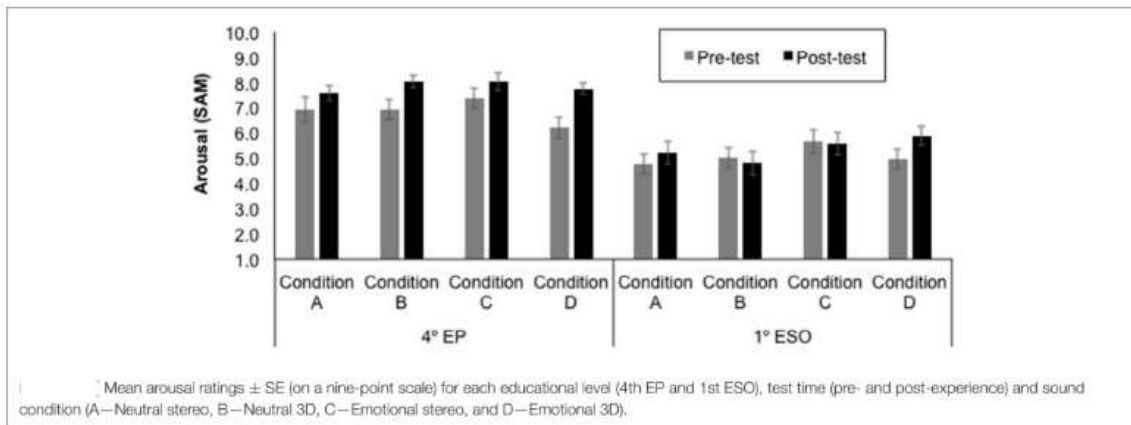


Figure 4 Mean Arousal

The above diagrams show the Statistical Mean Valence and Arousal self-commentary scores for each level of education, test time for processing (prior to and after the experience), and sound quality (Neutral stereo, Neutral 3D, Emotional stereo, and Emotional 3D). There were no significant differences in valence and arousal measurements based on educational level. However, the MANOVA analysis revealed several significant findings related to educational level, test time measurement, and their interaction. The majority of these students have shown higher overall positive valence ratings. Additionally, we observed a significant increase in valence ratings for both educational levels from the pre-test to the post-test. In addition, the participants in the 4th EP group showed higher ratings in the post-test compared to the pre-test. Further statistical analysis revealed that both groups experienced a shift in emotional valence from negative to positive during the post-test. The 4th DL group showed a significant increase in activeness. However, it is worth noting that there were no observed effects of sound tone or sound spatiality, nor any discernible relationship between the sound condition and the time of testing.

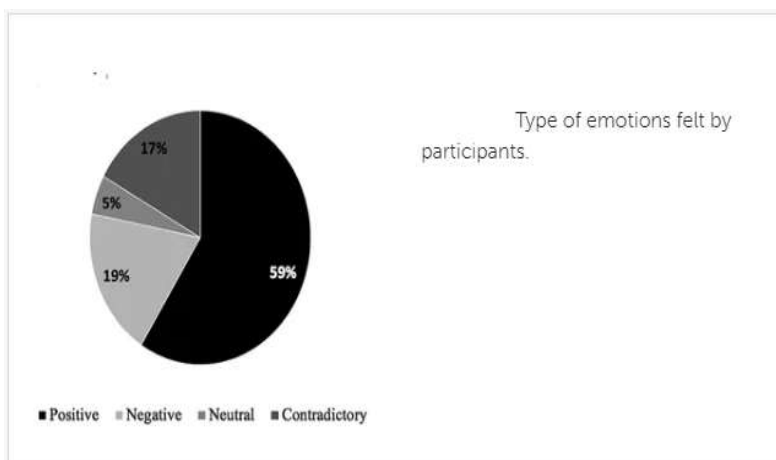


Figure 5 Emotions Distribution among Participants

The pie-chart provided displays the proportions of different emotional categories reported by the participants. Emotions of a positive nature are found to be prevalent in all of the conditions. It is quite intriguing that the division of the group into two did not result in any significant difference. The prevailing circumstances primarily affect the individuals, resulting in a 74% increase in feelings of happiness and sporadic occurrences of other emotions. In contrast, condition D elicited a greater range of emotions in participants, encompassing positive, negative, and conflicting types across the spectrum. This indicates that while condition D evoked a range of responses, including feelings of happiness, sadness, conflicting emotions, or even neutral ones, the other conditions only elicited either the specific emotion or both of the chosen emotions.

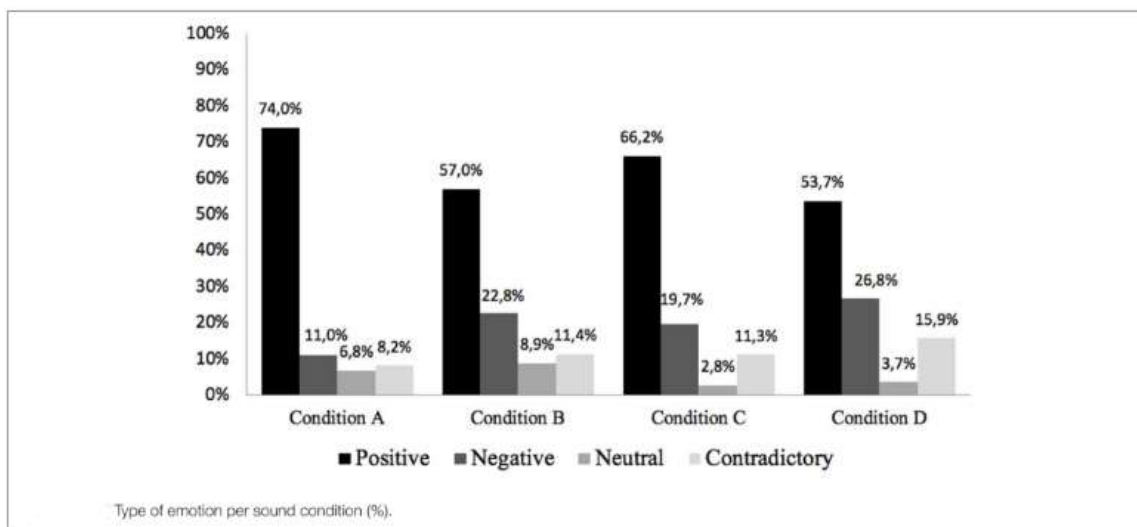


Figure 6 Emotion Per Sound

Subgroup Analysis

As for the study, subgroup analysis was employed to investigate the possible relationship of the effect of audio voice characteristics on the positive emotion reaction of audience in podcasting on different demographic variables. To identify the possible variations of emotion reactions in each age cohort, gender, educational level and cultural background, participants placed themselves into several categories.

Age

Participants were divided into three age groups: College Bound: 18-30, Age 30-50, and Over 50. Data was analyzed and identified the younger seasons (18-30 years) having more positive screen time with hosts who use voice modulation (mean rating of 4.8) as opposed to older groups. On the contrary, among the participants from age 50, they expressed a positive rating for the warm hues (average rating of 4.7) and as well moderate pacing (mean rating of 4.6).

Gender

Participants were divided reflecting the gender proportion by forming two groups of males and females. The females also demonstrated a stronger inclination toward podcasts with the hosts presenting moderate pitch range's (mean rating of 4.5 them). The males on the other hand only preferred the moderately pitch range podcasts (mean rating of 4.3 them). Furthermore, men displayed a more favorable include the participants of male sex in the positively emotional reaction to podcasts with average modulation (mean rate of 4.7) of female

participants (mean rate of 4.6).

Education Level

Participants were grouped based on their education level: first here, we have categories of people based on their education background which is high school or below, some college or bachelor's degree and postgraduate education. Education at post-graduate level has significantly increased positive emotionality towards podcasts featuring warmer tone of voice (mean rating of 4.7) that was perceived by individuals with low level of education (i.e. > 6.5 in scales of 0 to 10).

Cultural Background

The investigation into the impact of cultural ideas was performed by allocating participants in each group according to their culture. Findings based on cultural diversity show that participants with the diverse cultural backgrounds will have a much higher positive emotive response to podcasts with hosts incorporating multidynamic modulation (mean rating of 4.8) compared to those from the homogenous cultural backgrounds (mean rating of 4.6).

Additional Findings

In addition to the prominent findings about the impact of audio dialogue characteristics on audience favorable emotions during podcasting, several related studies have emerged. An analysis of user feedback revealed a correlation between the duration of a podcast and the retention rate of listeners until the finish. Episodes with duration of less than 30 minutes were highly correlated with increased happy feelings, while lengthier episodes, lasting between 30 and 59 minutes, did not exhibit significant differences. This demonstrates that a tightly constructed and clearly articulated material, which captivates the viewer, can generate a heightened sense of enjoyment and satisfaction.

Listeners exhibited recurring patterns and expressed a preference for podcasts with superior audio production quality. Research suggests that factors such as clear acoustics, reduced background noise, and appropriate sound levels per person have a beneficial impact and contribute to a positive emotional state. Podcasters should facilitate the acquisition of high-quality audio equipment and the development of expertise in editing procedures to ensure exceptional sound quality in their output.

Authenticity has now become a crucial measure of the effectiveness with which podcast storylines are able to capture and evoke emotional investment from listeners. A significant number of individuals conveyed that a host who wholeheartedly embodies the teachings of the gospel in their daily existence was the most reliable signal. In order to capture their interest, they believed that a host must be deeply engaged, genuine, and open in their narrative. Listeners were more likely to be emotionally affected by podcasts that, rather than being rigidly planned or dishonest, allowed for more authentic and straightforward sentiments. The relationship between establishing trust and delivering authentic material to the audience is quite significant.

A multitude of listeners provided feedback on the podcasts, expressing their distinct and diverse perspectives and viewpoints. Content that reflects diverse backgrounds and lifestyles is considered more appealing to a broader audience, resulting in a more varied and enjoyable listening experience. Inviting panelists from diverse backgrounds will engage a significant number of followers and expand the podcast's reach, making it more appealing to a wider audience.

Various genres of podcasts and styles of storytelling can evoke deep emotions in listeners. The audience can establish a stronger connection with content that revolves upon personal narratives, difficulties, and achievements, so eliciting a more profound emotional impact on the audience. Acquiring an understanding of the emotional fluctuations that elicit responses from audiences and predicting the themes that fulfill their desires

can result in devising strategies for content creation that are designed to enhance attendance and evoke good feelings.

Discussion

Interpretation of Results of Research Questions

Question 1: The analysis of the podcast revealed that the hosts' changing genders had a significant impact on the listeners' positive emotional response. Based on the data, it seems that the audience's emotional response can vary depending on factors such as pitch, tone, and voice modulation. In a similar vein, the gender of the podcast hosts appears to potentially influence the thoughts and level of emotional engagement experienced by listeners. The paper realized that male hosts provoked different emotional reflexes in comparison to female hosts, which underscored the importance of gender in podcasting.

Question 2: Developing an audio voice model which pleases the audience is a momentous goal achieved towards comprehending ways in which audio content may prompt emotional engagement. To help podcast hosts communicate with efficacy upholding audience satisfaction, this study highlights the key vocabulary of positive emotions.

Question 3: By assessing the emotional stance of the listeners pre- and post- listening to podcasts, the study gained insight into the prospective impact of video content on their feelings. The results underlined the importance of podcasts in evoking positive feelings and improving listeners moods and attitude. By evaluation the emotional reflexes of the audience pre- and post- listening, the team determined whether or not the goals were attained. This information allows the host to consequently make more conceptualized episodes and improve general satisfaction scores.

Question 4: The assessment of the link between gender features and positive voice models of podcast hosts unveils the intricate interdependencies amid audio features, audience perception, and personal gender identity. Results illustrate that the audience tends to depict vast preferences and emotional reflexes anchored on the gender of the hosts, especially when the latter's voice is in line with a give gender stereotype.

Comparison with Previous Studies and Theoretical Frameworks

After extensive research, this paper gained valuable insight into the effect of audio voice traits on the podcast industry and the relation formed amid listeners and the hosts. In Lindgren's thorough investigation of acclaimed podcasts in the UK and Australia, it becomes clear that the key to captivating listeners lies in the skillful combination of emotion and authentic storytelling [5]. Therefore, it aligns with research findings that highlight the power of smooth voices in evoking emotions and recognising the importance of vocal qualities in fostering audience engagement. In Zoghaib's study on voice influence, the profound impact of vocal states, such as intonation and pausing, on emotions and perception is highlighted, regardless of the podcast context[6]. Similar to conducting research, the study of voice interaction and artificial intelligence also explores the connection between speech and emotion. The way we communicate through our voice plays a significant role in expressing our character. This study explores the connection between audio voice attributes and the emotional response of podcast listeners.

Another crucial aspect is the podcasters' communication styles, which have a direct impact on the development of parasocial interactions between the podcasters and their audience, as highlighted by Tobin and [7]. The authors argue that first-person narratives significantly increase the listeners' level of engagement. In accordance with the literature review and reference theoretical framework, it is accurate to denote that audio voice features have a significant role in the context of podcasting, particularly in emotional engagement. By including insights garnered from already published articles this paper avails crucial insights into the effect of voice traits on the audience in the industry under scrutiny

Implications for Podcasting Practices and Content Creation

The studies mention within this brochure have availed valuable insight for competing firms and affords them guidance in the field, where efficacy is key. Additionally, the voice traits focus on involving the audience through narratives told from first-person view-points, prompting vast emotions. This denotes the essence of voice features in podcasting. During the creation of a podcast, it is essential for broadcasters

to keenly consider aspects such as emphasis of speech, speed, pitch, and tone to deliver the desired emotional stimulus to the audience with utmost efficacy.

Furthermore, the hosts' are keen on establishing authentic relations with the audience is seen in their endeavors to nurture relationships with a broader audience. By emulating a natural language approach, hosts can link with their audience on personal levels, sharing life experiences that are relevant to them. This nurtures a more vibrant attachment and involvement between the host and listeners. Essentially, the platform can help podcasters enhance their content production combinations, focus on storytelling, utilise voice modulation techniques, and strengthen audience connection.

Limitations of the Study and Suggestions for Future Research

This research has several limitations that warrant further investigation within the realm of scientific inquiries. At first, there are 50 participants (both males and females over the age of 18) which may not be enough given the wide range of preferences and demographic characteristics among podcast audiences. It would be beneficial to consider securing funding for larger and more diverse studies that can contribute to a broader understanding of the overall outcomes for a wider range of participants. After extensive discussion, the topic was thoroughly examined, with a focus on considering various factors such as the podcast environment and its content. It was found that individual listener preferences and previous experiences with this type of media greatly influenced emotional responses. It would be beneficial for researchers to also direct their attention towards the impact of individual differences on emotional engagement with podcasts. For example, one might examine personality traits, cultural backgrounds, and past experiences with similar content.

More emphasis is placed on the errors in audio voice features and their impact on the viewer's emotional responses, while other important factors such as pod content, format, and production quality are overlooked. For future studies, it may be beneficial to take a comprehensive approach that considers all aspects of interpersonal communication, such as vocals and other elements of podcasting. This will allow for a more detailed understanding of the level of audience attachment. However, the primary assessment relied on participants' self-reported emotional responses, which could potentially be influenced by biases and lack of honesty. Future studies could consider incorporating physiological measures or behavioral indicators alongside self-report data to enhance the analysis of strength scene processing.

Conclusion

Summary of Key Findings

Based on the findings, the study has provided important insights into how different types of tones in podcasting can elicit emotional responses from the audience. Through the utilisation of a sample size consisting of fifty individuals, equally divided between males and females, this study aims to explore the impact of positive emoticon-like voice models that mirror the attributes of podcasters in terms of gender on listener engagement. The investigation will employ various tests to examine this phenomenon. It is clear that vocal characteristics and audio voice are crucial in evoking emotional reflexes among listeners. Aspects like modulation, tone and pitch are equally important in this context. It is only right to mention that the participants depicted high levels of acceptance towards melodious voices which illustrated optimum modulation, unlike those that sounded monotonous and poorly modulated.

Theoretical and Practical Contributions

This short paper sought to discern the ways vocal qualities contribute to different forms of positive and negative emotions in different peoples by presenting a theoretical framework. The main message of the speech can be interpreted from the traits of voice timbre, frequency, and inflection in podcasts. The psychological enhancer of this type of content is shown in this way. In addition to that, gender norms should be raised to this discussion too along with media visual and the audience perception. As these issues can be a bit overlapping it might seem

more interesting to look at them together.

The research also focused on studying the emotional responses triggered by audio materials. Two different sound conditions were implemented to measure participants' electrodermal responses during sound exposures. Based on our initial findings, it is premature to make any definitive conclusions about the relationship between EDR and sound conditions. However, the apparent correlation between the emotional response to sound and the perception of sound location towards the end of the experiment may not actually indicate a true interaction. It was observed that in these situations, there were variations in emotional regulation among different populations, regardless of whether it involved a single stimulus or multiple stimuli. In certain conditions, there was a noticeable difference in the emotional reactions to "skate down" and "tac toe" scores. Furthermore, they experienced a range of positive emotions in all conditions, with condition D (a combination of an arousal sound design and 3D sound mix) offering the most variety.

This brochure chiefly provides academic advice for content creators (e.g. political analysts and podcasters) and independent journalists to follow. When establishing the structure for their content development, podcasters have to keep in mind factors that affect the listener's conduct. So, adding color to the genders and offering the slight nuances of them, the most powerful message for people of different age, culture and status could be delivered.

Final Thoughts and Recommendations for Podcast Producers

Generally, the data and results in this brochure suggest that vocal traits of audio settings are significant in impacting the feelings of the audience within the podcasting context. In terms of content generation in the podcasting sector, it is essential to identify the effect of voice nuances, like, pacing, tone, and pitch on engaging and captivating the audience. Thus, it is suggested that the hosts be keen on developing persuasive and engaging voice delivery strategies aligning with respective audience demographics. Moreover, the brochure has established that there are certain nuances concerned to gender which should not be ignored. It is crucial for the production crew to prioritize achieving voice spectrums and inclusiveness. Content developers can depict society by using varied elements and perspectives. This allows the audience to hear a variety of voices, encouraging togetherness and affirmation. The production team must also use storytelling techniques to evoke powerful emotions. Empathy, disclosure, and personification are crucial in nurturing a vibrant bond with the fan base and the clients, guaranteeing long-term involvement.

Concluding Remarks

This brochure has highlighted the essence of the voice characteristic in audio in the notion of emotional involvement while listening. By keenly curating these aspects, one gains insight into the chief elements that help evoke the feelings and better comprehend how these emotions are established in the subconscious of the audience. Having a grip on the findings in this paper, podcast hosts can improve their capability to engage listeners on a more intimate and impactful space by utilizing data and through workshops. With the industry rapidly evolving, podcasters need to embrace their discoveries and adopt them into the process of content creation. Highlighting the essence of various voices in storytelling, leaning towards the employability of authentic content, and constantly enhancing based on listeners feedback are important components in establishing vibrant relations with the listeners. Ultimately, the art of podcasters producing resonating and engaging content can evoke emotional reflexes that nurtures vibrant relations with the audience, consequently leading to the triumph of their media. Music has proven to be an asset in promoting the audiences' experience, nurturing involvement, and staying updated in the evolving world of podcasting.

References

1. Rime J, Pike C, Collins T. What Is a podcast? considering Innovations in Podcasting through the six-tensions Framework. *Convergence: The International Journal of Research into New Media Technologies* [Internet]. 2022 Jun 2;28(5):135485652211044. Available from: <https://journals.sagepub.com/doi/10.1177/13548565221104444>
2. RAJAR. MIDAS Measurement of Internet Delivered Audio Services SPRING 2020 [Internet]. 2020. Available from: https://www.rajar.co.uk/docs/news/MIDAS_Spring_2020.pdf9. Rime J, Pike C, Collins T. What Is a podcast? considering Innovations in Podcasting through the six-tensions Framework. *Convergence: The International Journal of Research into New Media Technologies* [Internet]. 2022 Jun 2;28(5):135485652211044. Available from: <https://journals.sagepub.com/doi/10.1177/13548565221104444>
3. Beniamini N (2020) The Infinite Dial 2020. Technical report. Edison Research.
4. Berry R. Part of the establishment: Reflecting on 10 years of podcasting as an audio medium. *Convergence: The International Journal of Research into New Media Technologies* [Internet]. 2016 Jul 8;22(6):661–71. Available from: <https://journals.sagepub.com/doi/10.1177/13548565166321059>. Rime J, Pike C, Collins T. What Is a podcast? considering Innovations in Podcasting through the six-tensions Framework. *Convergence: The International Journal of Research into New Media Technologies* [Internet]. 2022 Jun 2;28(5):135485652211044. Available from: <https://journals.sagepub.com/doi/10.1177/13548565221104444>
5. Lindgren M. Intimacy and Emotions in Podcast Journalism: A Study of Award-Winning Australian and British Podcasts. *Journalism Practice*. 2021 Jun 25;17(4):1–16.9. Rime J, Pike C, Collins T. What Is a podcast? considering Innovations in Podcasting through the six-tensions Framework. *Convergence: The International Journal of Research into New Media Technologies* [Internet]. 2022 Jun 2;28(5):135485652211044. Available from: <https://journals.sagepub.com/doi/10.1177/13548565221104444>
6. Zoghaib A. Persuasion of voices: The effects of a speaker's voice characteristics and gender on consumers' responses. *Recherche et Applications en Marketing (English Edition)*. 2019 Mar 11;34(3):83–110.
7. Tobin SJ, Guadagno RE. Why people listen: Motivations and outcomes of podcast listening. Rubinelli S, editor. *PLOS ONE* [Internet]. 2022 Apr 6;17(4). Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0265806>
8. Huang KL, Duan SF, Lyu X. Affective Voice Interaction and Artificial Intelligence: A Research Study on the Acoustic Features of Gender and the Emotional States of the PAD Model. *Frontiers in Psychology*. 2021 May 4;12.
9. Sienkiewicz M, Jaramillo DL. Podcasting, the intimate self, and the public sphere. *Popular Communication*. 2019 Oct 1;17(4):268–72.
10. Kreijns K, Xu K, Weidlich J. Social Presence: Conceptualization and Measurement. *Educational Psychology Review*. 2021 Jun 22;34.
11. Aghakhani N, Oh O, Gregg DG, Karimi J. Online Review Consistency Matters: An Elaboration Likelihood Model Perspective. *Information Systems Frontiers*. 2020 Jun 27;
12. Herrando C, Constantinides E. Emotional Contagion: A Brief Overview and Future Directions. *Frontiers in Psychology* [Internet]. 2021 Jul 16;12. Available from: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.712606/full>